DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

TAXING NEW FINANCIAL PRODUCTS: A CONCEPTUAL FRAMEWORK

Jeff Strnad
California Institute of Technology
University of Southern California

Forthcoming, Stanford Law Review, March 1994

SOCIAL SCIENCE WORKING PAPER 819

August 1993
Revised January 1994

TAXING NEW FINANCIAL PRODUCTS: A CONCEPTUAL FRAMEWORK

Jeff Strnad

Abstract

Two desirable properties for a tax system that must specify tax treatments for new financial instruments are consistency and universality. A tax system is universal if the system can designate a tax treatment for any cash flow pattern. Consistency requires that the tax treatment for each cash flow pattern be unique. A third property, linearity, holds if dividing the cash flow into different combinations of securities will not affect the tax treatment.

One way to achieve consistency and universality is to construct a tax system with a single systematic pattern of taxation, such as cash flow taxation or accretion taxation. But this extreme degree of homogeneity is not necessary. Consistent and universal tax systems can harbor radically different treatments for different types of transactions.

"Bifurcation approaches" divide a new financial instrument into certain prototype transactions with known tax treatments. The tax treatment for the new instrument is the sum of the tax treatments of the prototypes that sum up to the instrument. "Integration approaches" use rules that tax aggregates of instruments within the taxpayer's portfolio.

Bifurcation methods have a natural connection to linearity. These methods will not achieve consistency and universality in a nonlinear setting unless they are accompanied by elements of an integration approach. All universal and linear tax systems can be generated by "the spanning method," a specific kind of bifurcation.

Spanning method approaches are only a subclass of a broader set of integration approaches that achieve consistency and universality. In evaluating integration approaches, a key property is continuity, the requirement that tax treatments do not jump in response to small changes in any given portfolio. Continuity is a generalization of consistency. The existence of jumps leads to the possibility of serious tax manipulation of the same sort that would arise from inconsistencies.

The current U.S. tax system includes some direct inconsistencies. That is, the same transaction can be packaged different ways to achieve different tax results. These inconsistencies can only be eliminated by fundamental reform. Even the most powerful integration approaches cannot address the problem of direct inconsistencies.

This fact raises difficulties for authorities such as the Treasury Department and the courts who have only low level reform at their disposal. In promulgating regulations or deciding cases that involve new financial instruments, these authorities must choose rules using a second best approach. Loose ends in the form of inconsistencies or lack of universality are inevitable.

Table of Contents

Taxing New Financial Products: A Conceptual Framework

*Jeff Strnad**

The last two decades have witnessed the advent of ''financial engineering.'' Investment bankers, lawyers and other specialists have created innovative and sometimes complex financial instruments that allow investors and issuers to hedge risks, to speculate, and to achieve desirable tax results. The monetary volume of these new instruments is staggering.[1]

Financial innovation poses a deep challenge for tax policy. The current U.S. tax system works by using ''tax cubbyholes,'' a few idealized transactions for which the system specifies an exact tax treatment.[2] Since any given new

---

* John B. Milliken Professor of Taxation, University of Southern California, and Professor of Law and Economics, California Institute of Technology. I am grateful for valuable comments from participants in the Harvard Tax Policy Workshop, participants in the Southern California Tax Policy Group and, outside the workshop context, from Don Brown, Tom Griffith, Henry Hu, Louis Kaplow, Bill Klein, Mike Knoll, Ed McCaffery, Roberta Romano, Ted Sims and Al Warren. All errors are my own responsibility.

1. For several major categories of new financial assets, the aggregate principal amount of all outstanding notional principal contracts is in the trillions of dollars. The total notional amount for ''swaps,'' a set of recently developed instruments, exceeds the combined value of all shares listed on the New York and Tokyo stock exchanges. *See* Henry T. C. Hu, *Misunderstood Derivatives: The Causes of Informational Failure and the Promise of Regulatory Incrementalism*, 102 *Yale L.J.* 1457, 1548-60 (1993) [hereinafter, Hu, *Misunderstood Derivatives*]; Edward D. Kleinbard, *Equity Derivative Products: Financial Innovation's Newest Challenge to the Tax System*, 69 *Texas Law Review* 1319, 1320 (1991); William Glasgall & Bill Javetski, *Swap Fever: Big Money, Big Risks*, *Business Week*, June 1, 1992, at 102. The measurement of volume is in ''notional amount'' because many new instruments are ''derivatives'' with payouts that depend on the returns to a fixed amount of some other security or securities. *See* Hu, *Misunderstood Derivatives*, *supra*, at 1458-60, 1167 n.44. This fixed amount is the notional amount. For example, an instrument might pay an amount equal to the interest payments on $10,000,000 of a particular kind of floating rate debt.

It is not only the sheer volume of individual new instruments that is impressive. Financial innovators are introducing new instruments at unparalleled rates, and these instruments often reach large volumes after a very steep exponential increase. *See id.*

2. *See* Kleinbard, *supra* note 1, at 1320; Briffet, *Avoiding New Evils by Applying New Remedies: Taxes and the Cross-Border Transaction, J. App. Corp. Fin.*, Winter 1992, at 109, 110. Classification of new financial products in the face of rapid innovation is a problem for financial regulation and corporate law as well as tax law. In fact, the cubbyhole description put forward by both Briffet and Kleinbard had its origin in work by Henry Hu that focuses largely on these other fields of law. Professor Hu's work contains an admirable discussion of the classification problem in general. *See* Henry T. C. Hu, *Swaps, the Modern Process Financial Innovation and the Vulnerability of a Regulatory Paradigm*, 138 *U. Pa. L. Rev.* 333, 335-39, 392-412 (1989) (general discussion with application to bank regulation) [hereinafter, Hu, *Regulatory Paradigm*]; Henry T. C. Hu, *New Financial Products, the Modern Process of Financial Innovation, and the Puzzle of Shareholder Welfare*, 69 *Texas L. Rev.* 1273, 1292-1300, 1311-12

financial product is unlikely to fit squarely into one particular cubbyhole, the appropriate tax treatment for such products is often unclear.

This type of tax indeterminacy has fostered extensive debate and numerous proposals concerning the new financial product known as ''contingent debt.''[3] Traditional debt consists of an obligation specifying fixed interest and principal payments. Contingent debt, on the other hand, combines fixed payments with additional payments that depend on uncertain future events (such as the level of a commodity price or equity index).

The proposals put forward for taxing contingent debt reflect four theoretical approaches for taxing new financial products: bifurcation, integration, local pattern taxation and global pattern taxation.[4]

The ''bifurcation'' approach decomposes a new financial product into a collection of component instruments, each with a known tax treatment.[5] The 1991 version of the Proposed Treasury Regulations for contingent debt apply this

---

(1991) (classification problem with application to core principles of corporate law) [hereinafter, Hu, *Shareholder Welfare*]; Kleinbard, *supra* note 1, at 1320 n.1 and 1353 n.101; Briffet, *supra*, at 110 n.6.

3. *See, e.g.*, Randall K. Kau, *Carving Up Assets and Liabilities —— Integration or Bifurcation of Financial Products*, 63 TAXES 1003, at 1003-1014 (1990) (proposing an ''integration'' approach to minimize mismatching of timing, source and character); Edward D. Kleinbard, *Beyond Good and Evil Debt (And Debt Hedges): A Cost of Capital Allowance System*, 67 TAXES 943, at 943-61 (1989) (discounting the integration argument and traditional analyses in favor of a capital cost allowance system); David P. Hariton, *New Rules Bifurcating Contingent Debt —— A Mistake?*, 51 TAX NOTES 235, 235-39 (1991) (questioning the bifurcation approach advocated in the 1991 Proposed Regulations due to its inability to handle contingent debt instruments); Lokken, *New Rules Bifurcating Contingent Debt —— A Good Start*, 51 TAX NOTES 495, at 495-504 (1991) (countering Hariton's criticisms of the bifurcation approach); David P. Hariton, *More on Bifurcation of Contingent Debt*, 51 TAX NOTES 1075, 1075-76 (1991) (challenging the practicality of the bifurcation approach); Section of Taxation, The American Bar Association, *Report on Amendments to Proposed Regulation Section 1.1275—4: Proposed Regulations Regarding Certain Contingent Debt Instruments Under the Original Issue Discount Rules*, 53 TAX NOTES 1187, 1187-1204 (1991) [hereinafter *ABA Report*] (recommending simplifying rules even if at the expense of some coherence).

4. Indeed, the participants in the debate about the proper treatment of contingent debt apply their analyses directly to other financial instruments. *See, e.g.*, Kleinbard, *supra* note 1 (discussing the appropriate tax treatment for various equity swaps and other derivatives); David P. Hariton, *Equity Swaps, New Regulations, and Ed Kleinbard's Article*, 52 TAX NOTES 1221, at 1221-1224 (1991) (critiquing Kleinbard's integration argument in favor of a more rule-intensive approach).

5. Although ''bifurcation'' suggests division into *two* pieces, the literature uses the word to describe the decomposition of a financial instrument into two or more pieces. I follow that convention in this article.

approach.[6] Operationally, these Regulations call for subtracting the present value of the noncontingent portion (consisting of the fixed payments) from the issue price of the instrument. The noncontingent portion (fixed payments and their present value) is subject to the usual taxation rules for ordinary debt instruments. The remaining, contingent part of the instrument "will have the economic characteristics of one or more options or other property rights [which] ... can be taxed as they would be taxed if issued separately."[7] Thus, it may be necessary to divide the contingent part itself into separate pieces each having a known and distinct tax treatment.

"Integration," the logical complement of bifurcation, pools financial instruments together rather than splitting them apart. The resulting aggregate cash flow is taxed according to its "predominant characteristic."[8] For example, several commentators suggest that where a taxpayer fully hedges the contingent portion of contingent debt, tax policy should combine the hedge position with the contingent debt and treat the consequent riskless cash flow as ordinary debt.[9]

---

6. *See* Prop. Treas. Reg. 1.1275-4(g), 56 Fed. Reg. 8303 (1991). Treasury issued a later version of the Proposed Regulations for contingent debt on January 19, 1993 but then withdrew it along with other unpublished regulations on January 25, 1993 to allow reconsideration by incoming Clinton administration officials. This 1993 version abandons the bifurcation approach taken in the 1991 version. For a discussion of the 1993 version, see David C. Garlock, *A Primer on the New Proposed (Almost) Regulations for Contingent Debt Instruments*, 58 TAX NOTES 1225, at 1225-1230 (1993); David P. Hariton, *Contingent Debt: Putting the Pieces Together*, 58 TAX NOTES 1231, at 1231-1243 (1993).

7. Notice of Proposed Rulemaking FI-189-84, 1991-1 C.B. 835 (in the fourth paragraph of "Explanation of Provisions"). The pertinent text of the proposed regulation calls for treating the contingent payments "in accordance with their economic substance as payments pursuant to one or more options or other property rights." Prop. Treas. Reg. §1.1275-4(g)(4)(i), 56 Fed. Reg. 8308 (1991). Neither the explanation nor the text of the proposed regulation specify a method for choosing one particular decomposition of the contingent payments into other assets. More than one such decomposition may be possible.

8. Kau, *supra* note 3, at 1007.

9. *See, e.g.*, *ABA Report*, *supra* note 3, at 1199-1200; Kleinbard, *supra* note 3, at 953.

To see how hedging works, consider the following example. A company issues "gold bonds" that feature a payment that is tied to the price of gold on the date that the bond matures. The company also buys gold futures. Purchasing these futures in the right amount provides an exact offset against changes in liabilities on the bonds because of fluctuations in the spot price of gold. Thus, an increase in gold prices would result in a heavier obligation under the bonds and in equal and offsetting additional gains on the gold futures positions.

A possible motivation for the company to engage in this type of hedging is to achieve a lower cost of capital. Issuing gold bonds may be a cheap way to borrow money given inefficiencies in other segments of the capital market.

''Local pattern taxation'' applies a single generic treatment to all *new* financial products. This generic treatment includes rules for timing, characterization and source of cash flow. The term ''local'' emphasizes the fact that the generic treatment applies *only* to new financial products. The tax treatment of preexisting financial instruments may deviate sharply from that generic treatment. One should note that local pattern taxation may be combined with bifurcation and integration. For example, a recent American Bar Association report advocates decomposing contingent debt into contingent and noncontingent portions (as under bifurcation) but then taxing the contingent portion as a *single* unit under a set of generic rules.[10]

''Global pattern taxation'' applies a single generic treatment to *all* instruments. This approach directly responds to the idea that it is the current variety of distinct tax treatments for different existing investments that makes it especially difficult to prescribe tax rules for new financial products. For example, although equity investments with no current cash flows avoid taxation until realization occurs, interest is accrued and taxed on zero coupon bonds under the Original Issue Discount (''OID'') rules even though there is no cash flow from such bonds until maturity. When an instrument combines the features of a zero coupon bond and an equity investment that pays no dividends,[11] it is

---

This type of motivation appears to be very important in the real world. For example, the exploitation of capital market and regulatory inefficiencies is central to the classic ''arbitrage'' explanation for the growth of the multi-trillion dollar market in swaps. *See*, Hu, *Regulatory Paradigm*, *supra* note 2, at 350-53, 365.

10. *See ABA Report*, *supra* note 3, at 1195-1201. The report uses a ''cost recovery rule'' for timing, requiring that basis be recovered before any gain is recognized. Gain would be ordinary or capital depending on whether the transaction is part of the ''ordinary course of business'' or is an ''investment'' activity. The report does not specify sourcing rules but strongly urges the development of a uniform set of such rules. *Id.* at 1200.

A similar example is the ''expected value taxation'' system suggested by Professor Reed Shuldiner. *See* Reed Shuldiner, *A General Approach to the Taxation of Financial Instruments*, forthcoming 71 TEXAS L. REV. 243, 283-89 (1992). Under this system one would decompose each asset into a noncontingent portion consisting of all expected cash flows, and a residual contingent portion. An approach similar to the OID rules would determine the tax for the noncontingent portion, with income accruing based on the increase in present value as future expected cash flows approach. A realization approach would apply to the residual contingent portion, with no tax until the time when risk is resolved and final cash flows materialize. *Id.*, at 285.

11. For example, a ''debt'' instrument might consist of the right to receive $1000 plus the level of the S&P 500 index times $1 in ten years. The $1000 payment is fixed. Without the index component, the instrument would be a zero coupon bond. But the right to receive a payment contingent on the future value of the S&P 500 index is an equity interest equivalent to buying a stock that does not pay dividends and then selling the stock at a specified future time.

unclear whether income should be accrued and taxed or deferred until realization. Global pattern taxation would obviate such dilemmas by imposing a single consistent method (such as cash flow taxation or accretion taxation) to all instruments, including new financial products.

Assessing the relative merits of these disparate approaches requires a framework by which to compare them. The development of such a framework is one of the main purposes of this article. The analysis that follows relies heavily on two distinct desiderata for a good tax system: universality and consistency. Two variants of the consistency principle, linearity and continuity, also play an important role.

Universality requires that the tax system specify a tax treatment for every possible transaction. This principle is attractive both as an administrative goal and as an ideal in a system faced with financial innovation. If the tax treatment of particular portfolios of cash flow patterns is unclear, taxpayers and the govermnent will face heightened administrative costs. The government (in interaction with taxpayers) will have to specify rules for the ambiguous situations, and, prior to the development of such rules, taxpayers will be unable to predict the tax consequences of holding particular instruments or portfolios.

Part of the difficulty for the current U.S. tax system in dealing with financial innovation arises because the system is not universal. The tax treatment of novel transactions is sometimes unclear. But even if a tax system is universal, financial innovation poses another set of potential problems. Innovative packaging of a set of cash flows may result in a tax that differs from the tax that would be due if the cash flows were packaged in a more traditional manner. In a tax system where the same pattern of cash flows may have different tax consequences depending on the form chosen for transactions or portfolios, taxpayers will expend resources searching for the most advantageous form. At the same time, the government will be concerned that many tax treatments will become "elective" for taxpayers who can change these treatments by recasting their transactions or portfolios. Thus, even if the tax system is universal so that the tax treatment of every transaction is clear, substantial administrative costs may result if arrangements that are equivalent financially do not have the same tax consequences.

This problem motivates the idea of consistency. A tax system is *consistent* if and only if every cash flow pattern corresponds to a *unique* tax treatment. In such a system, it is not possible to manipulate tax outcomes by repackaging cash flows into different financial vehicles.[12]

12. Even commentators who believe that the consistency goal is unattainable still see it as normatively appealing. *See, e.g.*, Hariton, *supra* note 4, at 1224; Kau, *supra* note 3, at 1004.

Consistency is important not only because of the administrative costs that arise in its absence but also because of a close connection with "tax arbitrage." Tax arbitrage arises in its purest form when a series of transactions results in no net cash flow but provides tax advantages. Suppose, for example, that two portfolios result in the same cash flows but that the assets in portfolio one are capital assets while portfolio two produces ordinary gains and losses. Assuming both portfolios are likely to produce gains, an investor can make money at government expense (with high probability) by matching a long position in portfolio one and a short position in portfolio two. The long position results in capital gains while the short position creates ordinary losses in equal amounts, and the net cash flow from these two positions is zero. ''Conversion'' of ordinary income into capital gains occurs because ordinary income (such as wages) will be offset by the ordinary loss generated by the strategy and replaced by the capital gains from the long position. If capital gains rates are lower than ordinary income rates, the taxpayer receives a tax reduction without making any net investment.

This series of ''conversion'' transactions violates consistency because it is equivalent (in cash flow terms) to doing nothing, and the usual result for a taxpayer who does not engage in any transactions is that there are no tax consequences. In a consistent tax system, tax arbitrage is not possible. Since tax arbitrage tends to defeat distinctions set up in the tax laws (such as the distinction between capital gains and ordinary income) and tends to produce free money at government expense for well-capitalized taxpayers, the usual presumption is that tax arbitrage is an evil to be controlled. This view provides a normative basis for requiring consistency that goes beyond the goal of reducing administrative costs by making tax treatments determinate and unmanipulable.[13]

Bifurcation approaches appear capable of promoting the goal of consistency since these approaches tax each instrument an amount equal to the sum of the taxes on the components that make up the instrument. However, commentators have been extremely skeptical about the possibility of implementing operationally coherent bifurcation approaches. There are many ways to divide an instrument into pieces with known tax treatments, and different methods of division may result in different tax treatments for the instrument.[14]

---

13. There are a few cases where tax arbitrage can serve positive social goals. In these cases, arbitrage can induce price changes that are socially desirable. These price changes largely or entirely offset the tax advantages of the arbitrage itself. *See*, *e.g.*, Alan J. Auerbach, *Should Interest Deductions Be Limited?*, *in* UNEASY COMPROMISE: PROBLEMS OF A HYBRID INCOME-CONSUMPTION TAX 195 (Henry J. Aaron, et. al. eds, 1988) (tax arbitrage from leveraged holdings of tax-exempt securities may lower the price of borrowing for state and local governments).

14. Most commentators find this characteristic to be fatal. *See*, *e.g.*, *ABA Report*, *supra* note 3, at 1194-95; Hariton, *New Rules Bifurcating Contingent Debt —— A Mistake?*, *supra* note 3, at

Integration methods suffer from similar ambiguities. There is more than one way to aggregate sets of instruments into groups, and the overall tax results may depend on the particular choice of groupings. In addition, the proper way to characterize a particular aggregate of instruments may not be clear in a system replete with distinct and sometimes contradictory tax approaches.[15]

Taxation based on local patterns also entails potential consistency problems. New financial instruments may generate cash flows arbitrarily close to those of a preexisting instrument with a known tax treatment. Unless this tax treatment happens to correspond to the generic treatment for new financial products, instruments with nearly identical cash flows may incur very different tax liabilities.

Global pattern taxation is the only one of the four approaches that can achieve consistency and universality without an obvious operational or conceptual flaw. Because implementation of global pattern taxation would require systemic reform, this fact is of little comfort to administrators who must craft rules in a system arrayed with very different tax treatments that must be taken as given.

The problems with the four approaches have fostered significant frustration in the literature. Even prominent commentators who have developed and critiqued elaborate technical approaches have resigned themselves to relying on reform measures such as ''common law development'' or an ongoing dialogue between the Treasury Department and tax practitioners.[16]

Despite these bleak assessments, aspects of the theory of financial economics suggest that bifurcation and integration approaches are potentially viable. In particular, the literature on ''spanning'' provides an ''atomic theory'' of financial economics, in which it is possible to replicate the exact cash flow of *any* asset or portfolio by a ***unique*** combination of the elements from a specified set of assets.[17] The specified set of assets thus ''spans'' the universe of possible financial assets and is therefore called a ''spanning set.''

237; Kau, *supra* note 3, at 1004.

15. *See, e.g.*, *ABA Report*, *supra* note 3, at 1195; Hariton, *supra* note 4, at 1222.

16. *See, e.g.*, Hariton, *supra* note 4, at 1224 (ongoing dialogue); Lokken, *supra* note 3, at 504 (common law approach).

17. *See, e.g.*, Donald J. Brown, Charles D. Huijsmans, and Bernardus de Pagter, *Approximating Derivative Securities in f-Algebras*, in POSITIVE OPERATORS, RIESZ SPACES, AND ECONOMICS, edited by C.D. ALIPRANTIS, K.C. BORDER AND W.A.J. LUXEMBURG, at 171 (1992); Donald J. Brown and Stephen A. Ross, *Spanning, Valuation and Options*, 1 ECON. THEORY 3 (1991); Stephen A. Ross, *Options and Efficiency*, 90 Q. J. ECON. 75 (1976).

The possibility of spanning the entire economy with a fundamental set of assets raises the intriguing possibility that the tax law might simply specify the taxation of assets in the spanning set. To ascertain the tax treatment of any asset not in that set, including any new financial product, one would determine the unique decomposition of the asset into a weighted sum of spanning set assets and then add up the taxes for that weighted sum. Bifurcation *would be* a viable policy.[18]

Given the apparent simplicity of this method, one might ask whether it could work in a tax system where radically different approaches govern the tax treatment of particular transactions. For example, would the approach work if it is preordained that some transactions are subject to accretion taxation while cash flow taxation applies to other transactions? If the approach did work, it might be possible to retain some or all of the "tax cubbyholes" in current law while simultaneously taxing new financial instruments in a consistent manner.

Part I examines this question by studying the "spanning method," a bifurcation approach of the sort just outlined. Using the spanning method, it *is* possible to specify some "cubbyhole" tax treatments arbitrarily, and yet still construct a set of bifurcation rules that result in a consistent and universal tax system. However, the spanning method can succeed in achieving consistency and universality only for some configurations of cubbyholes and only if these cubbyholes are precisely defined. Short of fundamental reform, the present cubbyhole structure in the U.S. tax system precludes the successful use of the method.

Part II generalizes and extends the results in Part I by probing the relationship between the spanning method and the set of all consistent and universal tax systems. Part II begins by showing that for every consistent and universal tax system, there is an integration approach that can implement that system. In contrast, not every consistent and universal tax system reduces to the spanning method. This result highlights the fact that the categories of integration, bifurcation, and local pattern taxation are not mutually exclusive. The set of successful bifurcation and local pattern approaches is a subset of the set of successful integration methods.

The main task in Part II is to develop a logical taxonomy of theoretical approaches and then to relate this taxonomy to practical approaches such as bifurcation and integration. Developing the taxonomy and clarifying the

---

18. So would integration. It turns out that whenever there is a bifurcation approach that is consistent, there is also a consistent integration approach. However, the converse is not true. *See* text accompanying notes 72-77 *infra*.

relationship between the various practical approaches requires the use of two refinements of consistency, "linearity" and "continuity."[19]

Bifurcation approaches have a natural connection to tax systems that are *linear*. A tax system is linear when the tax on any transaction equals the sum of the taxes on any collection of subtransactions that comprise that transaction. Part II shows that a tax system is linear and universal if and only if it reduces to the spanning method. Thus, the spanning method is not just one particular kind of bifurcation but is the paradigmatic treatment for any linear tax system.

In nonlinear tax systems (such as the current system in the United States), the concept of *continuity* is important. Continuity exists when portfolios that are nearly identical have nearly identical tax outcomes. Continuity is a stronger condition than consistency but weaker than linearity. Thus, linear systems are a subset of continuous systems, which, in turn, are a subset of consistent systems. Part II shows that precisely the same goals that make consistency desirable, such as obviating problems of tax arbitrage, also make continuity desirable. In addition, since the current U.S. tax system has significant nonlinearities, certain continuous and universal integration approaches are the strongest candidates for dealing with new financial products. Unfortunately, despite the strength of these integration approaches compared to the spanning method, none of them are immediately applicable to the current U.S. tax system. Fundamental reform would have to precede their successful application.

Consequently, this article ultimately adds to the skepticism in the literature about bifurcation, integration and local pattern approaches.[20] Short of systemic reform, crafting rules for taxing new financial products requires difficult, "second best" choices. Nonetheless, the conceptual framework developed in the article clarifies the available approaches to taxing new financial products and may inform the ongoing debate about whether and how to institute such reform. In addition, the framework transcends the new financial products area. Bifurcation and integration techniques exist in many other areas of tax law. The framework developed in Part II adds insight whenever the law requires either that a

---

19. Consistency, continuity, linearity and universality are ideals of operational coherence for the tax system. However, this set of ideals is not comprehensive. There may be situations where it is desirable to sacrifice operational coherence for other goals such as economic efficiency or distributional equity.

Nonetheless, because coherence is a significant concern for both taxpayers and administrators, the tax structure implications of different operational coherence norms are important. Moreover, knowledge of the circumstances under which these norms are not fully attainable is valuable for courts, administrators and legislatures who must balance competing goals in designing and implementing tax rules.

20. *See* notes 14 and 15 *supra* and accompanying text.

transaction be divided into pieces and taxed according to the sum of the taxes on the pieces or that a series of transactions be aggregated and treated in a particular way independent of how the individual transactions would be taxed. Part III concludes by discussing the tax policy implications of the results in the article.

I. THE SPANNING METHOD

''Spanning'' studies from the finance literature provide a critical starting point for discussing necessary and sufficient conditions for tax systems to be consistent and universal. The ''spanning'' literature ranges from straightforward early work by Professor Stephen Ross[21] to recent work set in a very advanced mathematical context.[22] As mentioned above,[23] the spanning approach views any financial instrument as reducible to a combination of assets from a specified collection called the ''spanning set.'' Under the ''spanning method'' the tax on the income from an instrument is set equal to the sum of the taxes on the income from the spanning set assets that comprise the instrument.

In order to clarify the results from the spanning literature and to apply them to the taxation of new financial products, this Part proceeds in four sections. First, section A develops a simple example consisting of an economy with three financial assets to illustrate the operation of the spanning method. Second, section B employs this example to show that it is possible to use the spanning method to construct a consistent and universal system even when the tax law requires that particular existing assets be taxed in radically different ways. Third, section C shows that despite this positive result, consistency becomes unattainable when this system is ''overconstrained,'' that is, if the system has too many assets with predetermined tax treatments. In addition, even when accommodating disparate tax treatments is possible, consistency may require a considerable sacrifice in terms of ''tax aesthetics.'' The treatment of some transactions will conflict with most major conceptions of how such a transaction should be taxed. Finally, section D provides some policy perspectives on these problems, arguing that implementation of the spanning method remains infeasible absent fundamental reform of the current system.

---

21. *See* Ross, *supra* note 17.

22. *See* Brown, Huijsmans & de Pagter, *supra* note 17; Brown & Ross, *supra* note 17.

23. *See* text accompanying note 17 *supra*.

A. Spanning in a Simple Economy

Consider an economy that lasts only two one-year periods.[24] Within this economy there are three financial assets. Two of the assets are zero coupon bonds. One bond costs $100 initially (at "time 0") and yields $110 at the end of year one. The other bond costs $100 at time 0 and yields $121 at the end of year two. Thus, interest rates are 10 percent per year for both years in the model.

The final asset is a "stock," representing the right to collect a particular cash flow at the end of year two. The amount of the cash flow is uncertain, but it will take one of five possible values: $121, $242, $363, $484, or $605. Each of the five outcomes is equally likely, and no further information is available about the likelihood of any outcome until the end of the two years. Consequently, the stock's expected final value is $363.[25] Assuming that investors are risk neutral,[26] this $363 has a present value of $300 at time 0, which is therefore the price of the stock at that time. Furthermore, the stock will appreciate to $330 at the end of the first year.[27]

Finally, it is also convenient to assume that the five stock outcomes correspond to the five possible "states of the world." In other words, only five future environments are possible, and any "risky" endeavor is risky only because its outcome differs depending on which environment emerges.

The first step in applying the spanning method is to identify assets derived from the stock and the bonds that "complete the market." In economics

---

24. Two periods are the minimum needed to be able to distinguish between cash flow and accretion treatment. The difference between the two treatments is that cash flow treatment allows deferral of the tax on gains until realization. Consequently, studying deferral requires a model with at least two periods. If there is only a single period in which all cash flows occur, accretion and cash flow taxation will be identical.

25. $363 is simply the average value of the stock. Each of the five outcomes occurs with probability 1/5 so that the average outcome is $363 = (121 + 242 + 363 + 484 + 605) / 5$.

26. Assuming that investors are risk averse would not materially affect any of the results that follow. Those results depend only on the ability to replicate any possible cash flow pattern with some combination of assets from a spanning set. This ability is a function of the set of assets available and not of the degree of risk aversion of the populace.

27. Since investors are risk neutral, they will use the riskless rate of 10 percent to discount expected cash flows to present value. Since the initial uncertainty about the risky outcomes in the economy is not resolved at all until the end of year two, each risky asset will increase at a 10 percent rate until that time. Thus, during the first year the stock's price will increase by $30 over its initial value of $300.

terminology, a securities market is "complete" if for each state of the world, there is a portfolio that yields a positive amount in that state and zero in all others.[28] Each such portfolio provides "insurance" against a particular state occurring.[29] When portfolios may consist solely of combinations of the stock and the two bonds, it is impossible to construct an insurance portfolio for each state. Indeed, the only asset with varying returns is the stock, and it has positive returns in all five states. Completing the market, if it is possible at all, requires the creation of additional assets.

A central result in the "spanning" literature is that it is possible to complete the market if and only if there is some asset or portfolio that has distinct returns in each possible state of the world.[30] In that case, the ability to create call

---

28. The existence of complete markets simplifies the task of designing a tax system that is consistent and universal. In particular, the "spanning method" discussed in the next section works only when markets are complete. The spanning method is important since, as shown in Part II, a broad class of other approaches that guarantee consistency and universality reduce to that method. *See* text accompanying notes 73-76 *infra*. However, the existence of complete markets is not a sufficient condition for consistency and universality. These properties can only be jointly present if the tax system has a certain degree of coherence. This article shows that, absent comprehensive reform, the current tax system lacks that degree of coherence.

If the results were not negative, it would also be important to consider the possibility that markets are incomplete. Given incomplete markets, it would be possible to relax the universality condition by requiring a specified tax treatment only for every *attainable* cash flow combination. It would not be necessary to specify tax treatments for the "missing" financial instruments corresponding to cash flow outcomes that are not attainable. The tax authorities would never face the need to know how to tax these instruments since these instruments by definition would not be present in the world.

29. The amount paid for an asset with a positive return in one state and zero in all others is equivalent to an insurance premium. The positive yield on that asset if the particular state occurs corresponds to the payment of an insurance claim when an insured event happens.

If all states are insurable, it is easy to see why markets are called "complete." An individual can protect against *any* contingency that might occur. Since the ability to pass risks onto those most willing to bear them is an important economic mission for capital markets, completeness is a desirable property.

30. Professor Ross proved this result in a two-period world with a finite number of states similar to the one postulated here. *See* Ross, *supra* note 17, at 84-86 (Theorem 4 and related discussion). Others have since proven similar results using models with a continuum of states of the world and multiple periods. *See* Brown and Ross, *supra* note 17 (continuum of states); FRANÇOIS R. VELDE, ESSAYS IN THE THEORY AND HISTORY OF OPTIMAL FISCAL POLICY, 66-70 (1992) (unpublished Ph.D. dissertation, Stanford University) (multiperiod model with continuum of states). These more general results require advanced mathematics, a framework using "Riesz spaces."

The analysis in this article is limited to a much-simplified framework of a five-state, two-period economy. A more complex model would yield results substantially similar to those options on that asset or portfolio ensures a complete market. A call option consists of the right (but not the obligation) to buy a particular asset at a prespecified price.[31] That prespecified price is called the "exercise price" or "strike price."

The basic idea behind this complete market result is simple: By creating call options, it is possible to "divide up" an asset with distinct returns in every state (such as the stock in this case) into a series of "insurance" portfolios that yield a positive return in one state and zero in all other states.[32] This series of insurance assets is a "spanning set."

Working out an example based on the five-state economy described above helps to demonstrate how this method of completing the market works. As a first step, it is convenient to label the states by reference to the stock returns:

Table I: States and Stock Outcomes

State	Stock Outcome
A	121
B	242
C	363
D	484
E	605

obtained here, and the added mathematical complexity would not add significantly to the intuitive understanding available from the simple model.

31. An option is a contract between two parties. The *buyer* (also called the *holder*) of a call option has the right (but not the obligation) to buy a particular asset, such as a stock, at a particular price as specified in the option contract. The *seller* (also called the *writer*) of the option agrees to sell the specified asset at the specified price to the holder of the option if the holder exercises the option.

32. Alternatively, one can ensure complete markets by using only puts on the asset or portfolio. *See* Ross, *supra* note 17, at 82 (Theorem 2 and related discussion). Puts consist of the right to *sell* an asset at a particular price. Both puts and calls are necessary in some cases where the exercise prices of all options are restricted to be positive. *Id.* at 84 (Theorem 3 and related discussion).

Next, consider call options on the stock that the holder can only exercise when they expire at the end of year two.[33] The value of each such call option at that time will be equal to the difference between the price of the underlying stock and the contractually specified strike price, if that difference is positive. Otherwise, the option will be worthless.[34]

From this set of call options, one can create insurance portfolios that yield positive returns in one state but zero in all other states. This task is easy for state E. A call option on the stock with an exercise price of $484 will be worth $121 if state E occurs and zero if any other state occurs.[35]

Creating insurance portfolios for states A, B, C, and D requires combining call options with differing strike prices. For convenience, denote a call option with a strike price of $X as C(X) and denote the stock as S. The following table shows various options and their payoffs in each state:

33. Call options that can only be exercised on the expiration date are called "European" call options. In contrast, the holder of an "American" call option can exercise the option at any time between the date the option is written and the expiration date. European options suffice in the example developed here because *all* risk in the economy is resolved at the end of year two. Options exercisable only at that time will parse all the possible risky outcomes.

34. Suppose that the strike price of a European call option is $X. If the price of the stock is $(X+Y) at the time of expiration and Y is greater than zero, then it is optimal to exercise the option. Since the option allows one to buy at $X and then sell at $(X+Y), the option will be worth $Y at the time of expiration. Conversely, if Y is negative at the time of expiration (meaning that the stock price is below $X), then the holder will not exercise the option. In that case, the option expires and is worthless.

It is important to keep in mind that a call option does not *obligate* its holder to purchase at the strike price. A call option gives one the *right* to purchase at that price. One can choose not to exercise that right. For a more in-depth introduction to options, see RICHARD A. BREALEY & STUART C. MYERS, PRINCIPLES OF CORPORATE FINANCE 483-534 (4th ed. 1991).

35. This option gives the holder the right to buy to stock for $484 at the end of year two. If states A, B, C, or D occur, the holder will not exercise the option since the stock will be worth at most $484. However, if state E occurs, the price of the stock will be $605, and the option will be worth $121 (= $605 - $484).

Actually, any call option with an exercise price that is greater than or equal to $484 but less than $605 would suffice for insuring state E. Using an exercise price of $484 merely aids in the exposition of the example.

Table II: Option Payoffs in Various States					
State	Option				
	C(484)	C(363)	C(242)	C(121)	S = C(0)
A	0	0	0	0	121
B	0	0	0	121	242
C	0	0	121	242	363
D	0	121	242	363	484
E	121	242	363	484	605

As indicated in the table, the stock is equivalent to an option with an exercise price of zero. Denoting an insurance portfolio that yields \$1 in state Z and nothing in any other state as P(Z), the following equations represent the combinations of the stock and call options on the stock that generate an insurance portfolio for each state:[36]

$$P(A) = (1/121) \times [S - (2 \times C(121)) + C(242)]$$
$$P(B) = (1/121) \times [C(121) - (2 \times C(242)) + C(363)]$$
$$P(C) = (1/121) \times [C(242) - (2 \times C(363)) + C(484)]$$

---

36. It is easy to verify these relations. For example, to verify the second relation (the one for P(B)), consider the following table of payoffs:

Table of Payoffs				
State	C(121)	−2 x C(242)	C(363)	Aggregate Payoff (sum of the previous three columns)
A	0	0	0	0
B	121	0	0	121
C	242	−242	0	0
D	363	−484	121	0
E	484	−726	242	0

As the table indicates, combining one long position in each of C(121) and C(363) with two short positions in C(242) results in a portfolio that pays \$121 in state B and zero in all other states. Dividing this portfolio by 121 yields P(B), paying \$1 in state B and nothing in each other state.

P(D) = (1/121) x [C(363) − (2 x C(484))]
P(E) = (1/121) x [C(484)]

Negative signs in front of call options in the formulae denote the sale, rather than the purchase, of those options. Hence, appropriate long and short combinations of the stock and the five specified call options on the stock create insurance portfolios for each possible state. As a result, the market is complete in this example.

One can fully describe any asset with payoffs at the end of year two simply by specifying its payoffs in each of the five states. Furthermore, any such asset is a combination of the five portfolios in the set {P(A), P(B), P(C), P(D), P(E)}. For example, consider an asset that pays $2 in state B, $5 in state D, and nothing in the other three states. This asset is equivalent to a portfolio consisting of two P(B)s and five P(D)s. Moreover, since the P(Z)s represent combinations of the C(X)s, one can also describe the asset as a combination of the five underlying options in the set {C(0), C(121), C(242), C(363), C(484)}. Thus, any asset with payoffs at the end of year two is equivalent to some combination of these five basic options.

It is also true that any asset from either the set {P(A), P(B), P(C), P(D), P(E)} or the set {C(0), C(121), C(242), C(363), C(484)} will replicate the one-year payoff of any other asset or portfolio in the economy. In the model, all economic risk is resolved at the end of year two. Since investors are risk neutral, the value of *all* assets in the economy, including the stock, the bonds and the options will increase at the riskless rate of 10 percent per year during the first year. Thus, any of the assets in the two sets will replicate the returns on the zero coupon bond which pays $110 at the end of year one on an initial investment of $100.

As the above discussion indicates, it is possible to replicate any two-year cash flow pattern by a portfolio of assets either from the set {P(A), P(B), P(C), P(D), P(E)} or from the set {C(0), C(121), C(242), C(363), C(484)}. Since payoffs in this model occur *only* at the end of each of the two years,[37] some combination from each of these sets can replicate any financial asset in the economy. By definition, then, each of these sets is a *spanning set*. As such, each set is potentially useful for constructing a universal and consistent tax system. However, to ensure that the system specifies a unique tax outcome for each instrument, the spanning set cannot be "overspecified." In particular, it must be impossible to remove an element from the set and still have a spanning set

37. A more complicated model would permit payoffs at any time. Such a model would be mathematically more complex, but the results would be similar. *See* note 30 *supra*.

consisting of the remaining elements.[38] A spanning set with this property of irreducibility is called a *minimal* spanning set.

In fact, both of the above sets are irreducible. Irreducibility is obvious for the set {P(A), P(B), P(C), P(D), P(E)}. Each of the assets represents a two-year return of $1 in one of the five states of the world and zero in all of the other states. Removing any one asset from this set makes it impossible to replicate a nonzero return in one of the five states at the end of year two. Slightly more complex reasoning establishes the irreducibility of the set {C(0), C(121), C(242), C(363), C(484)}.[39]

---

38. If it were possible to remove an asset from a spanning set and still have a spanning set, then that redundant asset would be equivalent to a linear combination of assets in the "reduced" spanning set. One could substitute this combination of assets for the redundant asset in any combination from the "original" spanning set used to represent a financial instrument. Thus, there would be two different representations for any such instrument from that spanning set. The tax treatment of the instrument would be ambiguous unless the redundant asset had exactly the same tax consequences as the equivalent combination from the reduced spanning set. One may assign tax treatments arbitrarily to the assets in a minimal spanning set and not worry about the tax system becoming inconsistent. That freedom vanishes when the spanning set is not minimal.

39. The argument works as follows. For the set {C(0), C(121), C(242), C(363), C(484)} to be minimal, it must be the case that each element is essential for spanning to occur. Refer to the returns in Table II, and consider first the problem of creating an asset that pays off only in state A. For this task, C(0) is clearly essential since it is the *only* asset with a nonzero return in state A. The next step is to show that C(0) may be combined with other assets so that the combination yields zero return in state B without also yielding zero in state A. The only asset in the set that will accomplish this task is C(121) since it is the only asset besides C(0) that has a nonzero return in state B. The combination C(0) − (2 x C(121)) yields a positive return in state A and a zero return in state B. However, this portfolio has the side effect of negative returns in states C, D and E. Thus, one must add as asset or assets that makes the return in these three states zero while retaining the zero payoff in state B and a positive payoff in state A. The asset C(242) is essential for this task since it is the only asset besides C(0) and C(121) that has a nonzero return in state C. It turns out that C(0), C(121) and C(242) suffice to produce an asset with nonzero return in state A and zero returns in all other states.

To see that the remaining two assets, C(363) and C(484), are essential elements of the minimal set, consider the problem of designing a portfolio that pays off only in state C. Asset C(0) is the only asset with a nonzero return in state A. Since any portfolio including asset C(0) would yield a nonzero return in state A, C(0) cannot be one of the building blocks for the desired portfolio. Given that C(0) is ineligible, C(121) is as well since it is the only other asset with returns in state B.

It turns out that the remaining three assets, C(242), C(363), and C(484), are all necessary for creating the portfolio with returns only in state C for the same reasons that C(0), C(121) and C(242) were necessary to produce an portfolio with payoffs only in state A. In particular, C(242) is necessary because it is the only remaining asset (after excluding C(0) and C(121)) that has nonzero returns in state C. Consequently, C(363) is necessary to cancel out the returns from C(242) in state D, and C(484) is necessary to cancel the returns in state E from C(242) and

It is no coincidence that each of the two minimal spanning sets defined above contains five assets. The model assumes that five states of the world are possible at the end of year two. At least five distinct assets are necessary to capture the distinct outcomes in these five states.[40]

Given the concepts just developed, it is possible to state two principles that are useful for designing a tax system capable of dealing with new financial products:

> *Unique Representation Principle*: In an economy with a minimal spanning set, any collection of cash flows has a unique representation as a combination of assets in that set.
>
> *Nonunique Minimal Spanning Set Principle*: If a minimal spanning set exists, it generally is not unique.

Rigorously establishing the Unique Representation Principle requires some linear algebra. But the intuition behind the linear algebra is easy to understand. If there were two different combinations of assets from the minimal spanning set that generated the same cash flows, then subtracting the first combination from the second would yield a new combination with zero net cash flow. This new combination would include at least one asset with nonzero weight since the first and second combinations are different by assumption. Consequently, one of the assets in the minimal spanning set would have to be a combination of other assets in that set.[41] This relationship contradicts the assumption that the spanning set is

---

C(363).

40. Readers familiar with linear algebra will realize that the minimal spanning set in the example must have five elements. Because there are five states of the world, the returns in the second period form a five-dimensional vector space, with the return in each state representing one dimension. The returns from the assets in the set {P(A), P(B), P(C), P(D), P(E)} constitute an orthonormal basis for this vector space since each asset in that set returns $1 in one distinct state and $0 in the other four states. Any other basis, i.e., any other irreducible set of asset returns that spans the space, must also contain five elements. Furthermore, the assets in any such basis must be an invertible linear transformation of the assets in {P(A), P(B), P(C), P(D), P(E)}. As a result, the argument in note 39 *supra* that the set {C(0), C(121), C(242), C(363), C(484)} is a minimal spanning set reduces to showing that the returns from this set are an invertible linear transformation of the returns generated by the assets in the set {P(A), P(B), P(C), P(D), P(E)}. For a general discussion of the relationship between vector spaces, bases and spanning, see KENNETH HOFFMAN & RAY KUNZE, LINEAR ALGEBRA 28-49 (2nd ed. 1971).

41. For example, suppose that the minimal spanning set is {x, y, z} and that the following two combinations result in the same cash flow:

$3x + 5y + 2z$;

minimal. If combining certain assets in the set replicates the returns of another asset in the set, then that asset must be extraneous. A spanning set is minimal only if it has no extraneous assets.

The above five-state economy example illustrates the Nonunique Minimal Spanning Set Principle. There were at least two minimal spanning sets for the economy in that example. More generally, it is possible to show that there are an infinite number of possible minimal spanning sets in any economy where there is more than one distinct asset. This result stems from the fact that there are many ways to recombine assets from one particular spanning set to generate another such set. For instance, in the five-state example above, one could arbitrarily add *any* positive amount, α, of C(0) to each of the assets C(121), C(242), C(363), and C(484) in the minimal spanning set {C(0), C(121), C(242), C(363), C(484)} and still have a minimal spanning set. The new set {C(0), C(121) + αC(0), C(242) + αC(0), C(363) + αC(0), C(484) + αC(0)} transforms into the set of insurance portfolios, {P(A), P(B), P(C), P(D), P(E)}, by first subtracting αC(0) from the last four assets and then applying the transformation from {C(0), C(121), C(242), C(363), C(484)} to {P(A), P(B), P(C), P(D), P(E)} given above.[42] There are an infinite number of positive numbers α, and each choice of α yields a different minimal spanning set. Therefore, there are an infinite number of such sets.[43]

$$2x + 3y + 7z.$$

Subtracting the second combination from the first yields:

$$x + 2y - 5z.$$

This third combination results in zero net cash flow since it is the difference of two combinations yielding identical cash flows. As a result, it must be true that

$$x = 5z - 2y.$$

In other words, holding asset x is equivalent to holding a long position in five units of z combined with a short position in two units of y, and any portfolio formed from a combination of x, y, and z can be formed from a combination of y and z alone.

42. *See* text accompanying note 36 *supra*.

43. A rigorous proof that there generally are an infinite number of minimal spanning sets follows from the fact that any set of assets generating returns that are an invertible linear transformation of the returns from the assets in the set {P(A), P(B), P(C), P(D), P(E)} is a minimal spanning set. *See* note 40 *supra*. There are an infinite number of such transformations so long as the dimension of the return space is at least one. In the example, the dimension of the return space is five.

B. A Possibility Result

The two principles derived in the previous section suggest using the *spanning method* to design a consistent and universal tax system. This method consists of three steps. First, one chooses a particular minimal spanning set. Second, one specifies a tax treatment for each asset in that set and a rule for determining the tax treatment for any possible combination of such assets. Third, one imposes the rule that the tax due on cash flows from any asset not in the minimal spanning set is the tax that would be due on the unique combination of minimal spanning set elements that generate the same cash flow pattern as the asset.[44]

The Unique Representation Principle states that for any given minimal spanning set and any given attainable pattern of cash flows, there is a unique combination of assets in the chosen minimal spanning set that will generate the specified cash flow pattern. This principle guarantees that a tax system generated by the spanning method will be consistent and universal. It will be universal because any asset (and thus any attainable cash flow pattern) is equivalent to *some* combination of minimal spanning set assets. It will be consistent because the equivalent spanning set combination for any asset (or cash flow pattern) is unique and each spanning set combination has a known tax treatment.

The spanning method begins with the choice of a single minimal spanning set for implementing the method. A choice is necessary because the Nonunique Minimal Spanning Set Principle indicates that more than one minimal spanning set exists. It is possible to design a consistent and universal tax system that would apply different minimal spanning sets to different asset groups. However, choosing a single minimal spanning set makes it unnecessary to determine which minimal spanning set applies to any given cash flow pattern and makes it unnecessary to check for inconsistencies caused by applying multiple minimal spanning sets to sets of assets generating the same cash flow pattern.[45]

---

44. The second step in the spanning method, specifying the tax treatment for each asset in the minimal spanning set, is similar to the "cubbyhole" approach under current law which specifies the tax treatment for certain familiar transactions. However, the spanning method includes a determinate way of deciding how to tax an asset that does not have a specified treatment because it is not in the minimal spanning set. Under current law, there is no such determinate method for assets that do not fall into any existing cubbyhole. As result, it is difficult for current law to deal with new financial products.

45. There are potential benefits from using more than one minimal spanning set. Some cash flow patterns may be easier to decompose (into spanning set assets) using one minimal spanning set while other cash flow patterns may be easier to decompose using a different minimal spanning set.

To apply the spanning method in practice requires more detailed specification of the tax rules than simply setting out the three steps above. Of particular importance is the need to choose a method for determining the tax treatment of *combinations* of minimal spanning set elements. One simple choice for this method is a *linear* rule: The tax on a weighted sum of minimal spanning set assets is the weighted sum of the taxes on the assets using the same numerical weights. Thus, if stocks x and y are in the minimal spanning set, the tax on the combination of five shares of stock x and two shares of stock y will be five times the tax on a share of stock x plus two times the tax on a share of stock y. One might distinguish the spanning method that uses this linear rule for taxing combinations of minimal spanning set elements by calling it the *linear variant* of the spanning method. Because this article only considers this linear variant, however, no such special terminology is necessary.[46] Thus, in all of the discussion below, the term "spanning method" shall mean the linear variant of the spanning method rather than the more general class of all possible spanning methods.[47]

To illustrate the application of the spanning method, consider once again the five-state example delineated above. Suppose that the minimal spanning set chosen to apply the method is the set {C(0), C(121), C(242), C(363), C(484)}. Suppose also that someone invents a new financial asset called "D-insurance." One share of D-insurance yields $121 in state D at the end of year two but yields $0 if any of the other states (A, B, C, or E) occur.[48]

Having chosen the minimal spanning set, the next step is to specify the tax treatment of each of the five assets in that set. Suppose that the tax code calls for

---

46. Clearly, it is possible to employ nonlinear rules for computing the taxes on combinations of minimal spanning set elements. The linear rule is particularly interesting because of the natural connection, described below, between the linear variant of the spanning method and bifurcation approaches. *See* text accompanying notes 73-79 *infra*. The linear rule also is natural for those who are accustomed to the present U.S. tax system. In general, the tax on n units of an asset such as a stock or a bond is simply n times the tax on one unit of the asset. Finally, variants of the spanning method with more complicated rules fall under the general category of integration approaches. The second and third steps of the spanning method assign a unique tax treatment to every possible cash flow pattern. Later portions of the article show that the set of all possible ways to assign a unique tax treatment to each possible cash flow pattern is a particular set of integration approaches. *See* text accompanying note 72 *infra* (Entire Integration Principle).

47. This terminology is important because some of the results below concerning the "spanning method" are only true when the linear rule applies for computing the tax due on a combination of minimal spanning set elements. *See, e.g.*, text accompanying notes 73-74 *infra* and note 74 *infra* (Spanning Method Principle: "spanning method" results in a "linear" tax system).

48. The name "D-insurance" stems from the fact that holding the instrument provides insurance against state D occurring.

taxing these assets on a cash flow basis, but at unequal rates: a 40 percent rate for C(0), C(121), C(242), and C(363), but a 20 percent rate for C(484). Conceptually, this pattern is pattern is tantamount to permitting a favorable ''capital gain'' rate on the C(484) asset while taxing the remaining assets at a higher, ''ordinary income'' rate.

D-insurance is equivalent to 121 units of the asset denoted P(D) above.[49] The unique decomposition of P(D) in terms of minimal spanning set assets is:

$$P(D) = (1/121) \times [C(363) - (2 \times C(484))]$$

The unique decomposition of D-insurance is therefore:

$$\text{D-insurance} = C(363) - (2 \times C(484))$$

Thus, under the spanning method, the tax on D-insurance is equivalent to the tax on one long position in C(363) plus the tax on two short positions in C(484). Applying the appropriate rates, the taxes faced by an owner of D-insurance at the end of year two in the various states are as follows:

Table III

State	Return on C(363)	Tax on C(363)	Return on −2C(484)	Tax on −2C(484)	Return on D-insurance	Tax on D-insurance
A	0	0	0	0	0	0
B	0	0	0	0	0	0
C	0	0	0	0	0	0
D	121	48.40	0	0	121	48.40
E	242	96.80	−242	−48.40	0	48.40

This analysis yields one peculiar result: in state E owners of D-insurance pay a tax of $48.40 even though D-insurance returns nothing in that state. This anomaly exists because the tax system applies different rates to cash flows from C(363) and C(484). Recall that the tax system treats D-insurance *as if* it were a portfolio of these two assets. Thus, when state E occurs, the $96.80 tax on the

49. *See* text accompanying note 36 *supra.*

gain from the long position in C(363) exceeds the $48.40 deduction on the capital loss from the two short positions in C(484).

Thus far the analysis has had an optimistic tenor. If the government can identify a minimal spanning set, it can fashion a tax system that is consistent and universal. This "possibility result" holds true regardless of how the government specifies the tax treatment for assets in the minimal spanning set. The treatment of different assets may be conceptually distinct.

Nonetheless, the positive tax on a zero net cash flow outcome for state E in the five-state example is disturbing.[50] This result violates the "obvious" precept that an investor should not pay any tax if there is no "income" or "cash flow." The cause of the result was the dissimilar treatment of certain assets. In particular, the system applied a more favorable tax rate to C(484) than to other assets. Equalizing the applicable rates on all assets would eliminate the problem. However, disparities such as applying a special tax rate to C(484) in the example may have countervailing theoretical or economic justifications. Although the spanning method allows the tax system to remain consistent and universal in the face of some desired disparities, there is a potential cost in terms of "tax aesthetics." The tax treatment of some transactions will not make sense under standard tax concepts.

---

50. One cannot justify this $48.40 tax on a zero net cash flow as an offset for some tax benefit resulting from the purchase of the D-insurance. To demonstrate this point, recall that D-insurance is equivalent to one long position in C(363) combined with two short positions in C(484). Assuming risk neutrality and an interest rate of 10 percent, one can calculate the prices of each option. A C(363) option pays $121 in state D and $242 in state E and nothing in any other state. The $121 and $242 cash flows have present values of $100 and $200 ($121 / $(1.10)^2$ and $242 / $(1.10)^2$ respectively). Given that states D and E each have a 1/5 chance of occurring, a risk neutral investor would compute the present value of C(363) as the sum of 1/5 of $100 and 1/5 $200, for a total of $60. Thus, the market price of C(363) must equal this $60 present value. A similar computation yields $20 as the price of C(484). As a result, the pre-tax price of D-insurance at time zero is $20, the $60 cost of buying one unit of C(363) less two times the $20 revenue from writing a C(484) option.

What are the possible tax benefits associated with purchasing D-insurance? The example in the text presumes that a cash flow tax system is in effect. In such a tax system, there is a deduction for net investment. At a 40 percent tax rate, the tax benefit for the $20 net investment would be $8. This amount would increase to $16 under an approach that considers the components of D-insurance separately: The holder would deduct the $60 cost of the C(363) against a 40 percent rate and would pay a tax at the special 20% rate that applies to C(484) on the receipt of the $40 from writing two C(484) options. In contrast to the possible initial tax benefit of $8 or $16, the tax at the end of year two is $48.40, and the present value of this tax as of the time of the initial benefit is $40. The resulting net tax of $24 to $32 in present value terms is totally inappropriate (under conventional tax reasoning) given that in pre-tax terms the net result for the investor is a *loss* equal to the entire initial investment of $20.

In the "real world," tax systems often contain disparities and the U.S. tax system is no exception. An accretion type of tax applies to some assets while tax is levied on others only upon realization of gains or losses. Corporate assets are subject to an extra layer of taxation. Different rates apply to income from "capital assets" and "ordinary assets." The loss carryover rules create an asymmetry between the taxes of losses and gains. The next section analyzes how these disparities affect the ability to apply the spanning method successfully. Unfortunately, some of the disparities cause "impossibility results" to replace the "possibility result" developed above. This consequence goes far beyond the problem that attaining consistency and universality using the spanning method may require a sacrifice in terms of "tax aesthetics."

C. Some Impossibility Results

Any proposal for taxing new financial products faces the real world constraint that numerous assets in the economy have a "predetermined" tax treatment, that is, a treatment that a policymaker must take as given in devising rules. Predetermined treatments are especially troubling when they are based on differing principles (*e.g.*, an accretion approach versus a realization approach). Such discrepancies preclude the use of global pattern taxation where a single principle (*e.g.* accretion) dictates the treatment of all transactions. The Treasury Department and the courts face precisely this scenario when determining the tax treatment of specific new financial products. Comprehensive reform to achieve a single global pattern for the taxation of all assets cannot be accomplished through regulations or court decisions.

The effect of being constrained by predetermined tax treatments is evident in the five-state example from the previous section. In that example, the available assets were two zero coupon bonds and a "stock." Under current U.S. tax law, a zero coupon bond is subject to "OID rules" that have an accretion aspect. The OID rules impute a stream of interest payments to the bond and impose a tax on these payments even though they are fictional.[51] On the other hand, current law

51. Pure accretion treatment would require taxing the total change in value of the bond each year rather than merely the "interest component" estimated under the OID rules. These two approaches differ because the total change in value of a bond during each year may differ from the amount of interest imputed to the bond based on the initial fixed schedule of payments. For example, if interest rates increase during the life of the bond, the capital value of the bond will decline so that the total gain in some years will be less than the amount of interest imputed to the bond. Nonetheless, the OID rules are a step toward pure accretion treatment. One can view the OID component as an estimate of the annual change in value of the bond, and this estimated increase is taxed even though it is unrealized.

generally does not impose a tax on the returns from "stock" until they are realized as distributions or as capital gains upon sale.

Taking these current law treatments as predetermined constrains the choice of tax treatments for the five assets in the particular minimal spanning set chosen for applying the spanning method. There is one specific combination of these assets that is equivalent to the stock and two other specific combinations that are equivalent to the bonds. The tax on these combinations must correspond exactly to the predetermined tax treatment of the stock and the bonds.

Suppose that the minimal spanning set chosen for applying the spanning method is {C(0), C(121), C(242), C(363), C(484)}. This set is a particularly instructive choice because all of the assets in the set are call options, instruments that have a specific tax treatment under current law. For now, however, assume that the tax treatment of these instruments need not be the same as under current law.

The stock in the hypothetical economy is simply the asset C(0), which is already an element of the minimal spanning set.[52] The initial value of the stock is 300, and there is no dividend or other realization event until the end of year two. At the end of year two, the shareholder receives a cash distribution (representing the total return on the stock) in exchange for each share. Thus, there is no tax at the end of year one, but the cash distribution net of the $300 cost per share is taxable income at the end of year two.

The two-year zero coupon bond appreciates from $100 to $110 during year one and then from $110 to $121 during year two. Under existing law, this appreciation results in OID income of $10 at the end of year one and $11 at the end of year two. Assuming a 40 percent tax rate, the bondholder will pay taxes of $4.00 and $4.40 at the end of years one and two respectively.

None of the five assets in the minimal spanning set {C(0), C(121), C(242), C(363), C(484)} corresponds exactly to the two-year zero coupon bond. However, the bond is a simple combination of two of the assets in that set, one of which is the stock. Denoting the bond as B(2) and the stock as S, the following identity applies:

$$B(2) = C(0) - C(121) = S - C(121).$$

52. A call option with a zero exercise price on a limited liability asset (such as the stock in the example) is identical (in monetary returns) to the asset. Limited liability implies that the lowest possible payoff is zero, and a call option with exercise price zero is the right to receive any return above zero.

Because the tax treatments of both the stock and the two-year zero coupon bond are predetermined, this equation implicitly specifies a tax treatment for C(121). The tax on C(121) must equal the tax on a position that is long one share of the stock and short one bond:[53]

$$C(121) = S - B(2).$$

A party holding C(121) would deduct the \$10 imputed interest on the bond at the end of year one, resulting in a \$4 tax benefit given a 40 percent rate. At the end of year two, the holder would deduct \$11 for that year's imputed interest on the

53. This equation is the "put-call parity equation" for a case where the put has zero value. The general form of the put-call parity equation is:

$$S + P = C + PV(X)$$

where S is an underlying asset, C is a call on S at exercise price X, P is a put on S at exercise price X, and PV(X) is the discounted present value of X, calculated using the time until expiration and a riskless discount rate. No put term appears in the equation in the text since the exercise price is \$121, the lowest possible outcome for the stock. A put will only be valuable if there is a possibility that the value of the underlying asset will fall below the exercise price. The bond ensures a return of \$121 at the end of year two. As a result, the present value of the bond is the present value of the \$121 exercise price.

bond.[54] In addition, the cash distribution paid in exchange for the stock at the end of year two would trigger a tax.

Because the tax treatments of the bond and the stock are predetermined, the policymaker no longer has the freedom to choose the tax treatment of C(0) and C(121), two of the five assets in the minimal spanning set. However, complete freedom to choose the tax treatment of the other three assets remains.

---

54. Under current law, corporate issuers may deduct the imputed interest payments on zero coupon bonds. *See* I.R.C. §163(e). However, it is unclear whether a deduction for imputed interest is available if a non-issuing entity (such as an individual investor) sells a zero coupon bond short.

Selling short involves borrowing the security that is sold. Consequently, the short seller must pay the lender any dividend or interest payments due on the security during the period that the seller holds the short position. When these payments are in *cash*, it is clear from the statute that *any* short seller may deduct the payments as "investment interest" so long as the security sold short is not a tax-exempt security. *See* I.R.C. §163(d)(3)(C) and §265(a)(5) (1994). The terms of the statute also include amounts "incurred" or "accrued," see *id.*, but it is not clear that these terms cover imputed interest since the short seller does not owe that interest to anyone.

A zero coupon bond will tend to appreciate in an amount equal to the imputed interest, thereby creating a loss in that amount for the short seller. In the absence of being able to deduct the imputed interest payments directly, the short seller will be able to deduct them as a loss (to the extent they are reflected in appreciation of the bond) upon closing the short position. This result is disadvantageous for the short seller (compared to deducting each year's imputed interest payments against ordinary income) since the loss will often be a capital loss and since the seller cannot deduct the loss until the time of sale.

A 1972 Revenue Ruling suggests that the government will treat obligations of a short seller other than cash payments discharging dividend or interest obligations as "amounts paid for replacing a borrowed security" which are nondeductible capital expenditures. *See* Rev. Rul. 72-521 (holding that short seller may deduct payments covering cash dividends on the stock but may not deduct either the payment of a nontaxable liquidating dividend on same stock or the cost of covering additional shares from a nontaxable stock dividend). However, the Revenue Ruling is factually distinguishable since it disallowed deductions for *nontaxable* dividends. The imputed interest from zero coupon bonds is taxable.

In general, to achieve consistency, a tax system must treat opposite positions symmetrically. In other words, the tax treatment that applies to a short position must be the "negative" of tax treatment for the corresponding long position. Otherwise, combining long and short positions would result in a net tax effect even though the cash flow outcome is equivalent to no position at all. Although special treatment (such as zeroing out the tax results) might be accorded to any situation where a short and a long position offset each other, such special treatment might not be very effective. A taxpayer could avoid the required special treatment by constructing positions that fall short of, but come close to, perfectly offsetting short and long positions. For a more complete discussion of this "approximate arbitrage" maneuver, see text accompanying notes 83-86 *infra*.

Conveniently, the treatment of these three remaining assets can be chosen to create a generic treatment for call options. In particular, the government could tax each call option as if it were equivalent to holding the underlying asset and borrowing an amount equal to the difference between that asset's value and the cost of the option at the time of purchase.[55] This generic rule is consistent with the predetermined treatments for C(0) and C(121). These predetermined treatments impute no borrowing for C(0) and $100 of borrowing for C(121). C(0) is the underlying asset, the stock. Consequently, the appropriate amount of imputed borrowing for C(0) under the generic rule is zero. C(121) has an initial value of $200 compared to the $300 initial value of the underlying asset, C(0). As a result, $100 of imputed borrowing is appropriate under the generic rule.[56]

This generic rule for call options conflicts with current law since current law does not impute borrowing to the holder.[57] However, given the predetermined treatments of the stock and the two-year zero coupon bond, consistency requires that an imputed borrowing rule apply at least to the C(121) call option.

One response to this problem is to revise the current rules for taxation of options. This reform would be very limited in scope, involving only three

---

55. The writer of a call option would receive symmetric treatment. Symmetric treatment of short and long positions is necessary to ensure consistency. *See* note 54 *supra*.

56. This generic treatment does not accord with the most general form of the put-call parity equation. That equation shows that the value of a call is a function of the value of the asset plus a loan *plus the value of a put*. The put has the same exercise price as the call and the loan represents the obligation to pay the exercise price on the expiration date of the option.

Although the suggested treatment in the text neglects the put, including the put value in the formula does not present a consistency problem. Because the stock yields a minimum of $121, the value of puts at strike prices of $0 and $121 (corresponding to the calls C(0) and C(121)) would be zero, and these puts would not result in any cash flows. *See* note 53 *supra*. Consistency requires there be no tax consequences from any such "investment." *See* text accompanying notes 12-13 *supra* (arbitrage transactions that cost nothing and yield no net cash flows should not have any tax consequences). *Cf.* text accompanying notes 49-50 *supra* (consistency may require tax on a zero cash flow outcome for an investment that is costly and that produces nonzero cash flows in some states of the world). As a result adding these worthless puts to any portfolio will not affect the sum of the tax treatments of the assets from the portfolio.

57. Sale of an equity option before it expires results in a capital gain or capital loss. The capital gain or loss will be short or long term depending on the holding period for the option. If the option expires worthless, the tax laws treat the expiration as equivalent to the sale of the option on the expiration date for $0. Exercising a call option results in the addition of the option price to the price paid for the stock for purposes of determining the total cost basis of the stock. *See* I.R.C. §1234 (1994).

Special mark-to-market rules apply to options on futures and other nonequity options. *See* I.R.C. §1256 (1994).

sections in the Code.[58] In addition, commentators have argued that the current treatment of options impedes the effort to design rules for taxing new financial products and has little or no justification on tax policy or tax theory grounds.[59] Reform of the option rules therefore could be regarded as an easy step that would be quite worthwhile if it led to a coherent way to tax new financial products.

Unfortunately, reforming the tax treatment of options will not suffice to ensure successful implementation of the spanning method. Other ambiguities and tensions deeply embedded in current tax law are substantial impediments. One example is the distinction between "stock" and "bonds." This distinction is so ambiguous that direct inconsistencies in taxation result.

Suppose, for example, that a corporation engages in a project with returns of 242, 363, 484, 605 and 726 in states A, B, C, D, and E, respectively. In terms of the economy in the five-state example, this project generates returns equal to the returns from holding one share of the "stock" and one two-year zero coupon bond. Consistency would require taxing the two-year coupon bond portion of the project under an OID approach. However, under current law, if the corporation finances the project with equity, it can defer the tax on all of the project's returns until the end of year two. In short, current law prescribes sharply different treatments for "stocks" and "bonds" without clearly distinguishing the two. Indeed, even the paradigmatic "stock" in the example contains a built-in bond component.[60]

The spanning method cannot deal with ambiguity that runs this deep. If tax policy requires distinct treatments for "stocks" and "bonds," these categories must be defined much more precisely. Implementing precise definitions, however, would require fundamental and comprehensive reform, overturning familiar tax treatments for numerous financial instruments.[61]

Professor Reed Shuldiner's rule of "expected value taxation" provides an example of this kind of reform. This rule divides financial instruments into a noncontingent portion representing the expected return from the instrument and a

---

58. I.R.C. §§ 1234, 1234A & 1256 (1994).

59. *See, e.g.*, Kleinbard, *supra* note 3, at 951; Hariton, *New Rules Bifurcating Contingent Debt —— A Mistake?*, *supra* note 3, at 239.

60. The stock returns at least $121 in all five states of the world.

61. Rules that rely heavily on the fragile distinction between equity and debt also create major problems in corporate law. *See* Hu, *Shareholder Welfare*, *supra* note 2, at 1286-1300 (whether a security is equity or debt or a known hybrid of equity and debt makes a big difference in the legal rights of the holder; financial innovation has produced many instruments that are hard to categorize, even given established rules for familiar debt/equity hybrids such as preferred stock and convertible bonds).

contingent portion representing deviations from that expected return.[62] The rule would treat the ''stock'' in the five-state example as the sum of three two-year zero coupon bonds and a risky residue with zero initial value:[63]

Table IV: The ''Stock'' and its Contingent and Noncontingent Components			
State	Return of Instrument or Component		
	Stock	Noncontingent Component	Contingent Component
A	121	363	−242
B	242	363	−121
C	363	363	0
D	484	363	121
E	605	363	242

This approach eliminates ''bond-stock'' ambiguities by providing a determinate way to isolate the ''bond'' component from any risky investment. However, applying this method across the board would require substantially revising the current treatment of equity instruments.[64]

62. *See* note 10 *supra* (summary of rule).

David Hariton has proposed an approach for taxing contingent debt obligations that resembles Professor Shuldiner's rule. Under Hariton's scheme, taxable income in the form of interest or OID would accrue on the revised issue price of the debt at no less than the applicable annual federal rate. *See* Hariton, *New Rules Bifurcating Contingent Debt —— A Mistake?*, *supra* note 3, at 238.

63. Three two-year zero coupon bonds would cost $300 at time zero and yield $363 in all states of the world at the end of year two. Since the stock's value at time zero is $300, the value of the ''contingent'' portion of the stock that remains after subtracting the three bonds must be zero. Table IV also illustrates why the contingent portion has zero value. The expected return of the contingent portion is zero, and risk neutral investors will value an asset with zero expected return at $0.

64. In addition, applying this approach (or another approach with a similar ability to clarify the bond/stock distinction) only to *new* financial instruments will not result in consistency. The inconsistency problem arises from the tax treatment of *old* financial instruments. Two or more bond/stock combinations with identical cash flow patterns incur different tax liabilities under the existing rules.

D. A Perspective on the Results

The previous sections demonstrate that a tax system can be consistent and universal even when it treats certain classes of transactions quite differently. For example, one type of asset might be subject to an accretion rule while another is subject to a realization rule.

However, the flexibility in choosing tax treatments is not unlimited. If the rules "overconstrain" the system by specifying a tax treatment for too many assets, the spanning method will fail. In particular, the number of distinct predetermined tax treatments cannot exceed the number of states.[65] In addition, even if there are "many" states (as there undoubtedly are in the real world), the law must not contain direct inconsistencies such as that between the current tax treatment of options and the current tax treatment of stocks and bonds.[66] Finally, even if the system is not overconstrained and does not contain direct inconsistencies, choosing radically different theoretical approaches for certain paradigmatic transactions can prove costly in terms of tax system aesthetics: Consistency may require a tax treatment for some transactions that does not make sense under *any* theoretical or conceptual approach. For example, investors may have to pay a significant tax in some instances where there is zero net income and zero net cash flow.[67]

Unfortunately, reform of the current system would require much more than making some simple adjustments and accepting some unpleasant aesthetics. The current system contains major direct inconsistencies. Although it would be easy to remove some inconsistencies, such as those stemming from the tax treatment of options, only comprehensive reform could remove others (such as those stemming

---

Furthermore, applying a generic treatment to new financial instruments may introduce new inconsistencies into the tax system. One may be able to invent new instruments that replicate the returns of the "bond" and "stock" assets to which the "old" rules apply. For example, an appropriate mixture of assets from the set {P(A), P(B), P(C), P(D), P(E)} consisting of insurance portfolios can replicate the payoffs of the stock:

$$S = 121\ P(A) + 242\ P(B) + 363\ P(C) + 484\ P(D) + 605\ P(E).$$

Unless the generic treatment for this mixture of insurance portfolios is the same as the treatment of the stock under the existing rules, inconsistency will result.

65. Thus, in the example developed in the text, having five states permits arbitrary specification of tax treatments for *at most* five instruments without introducing inconsistency. *See* text accompanying notes 55-56 *supra*.

66. *See* note 38 and text accompanying notes 52-58 *supra*.

67. *See* text accompanying notes 49-50 *supra*.

from the lack of a clear delineation between "stocks" and "bonds").[68] A quick survey of current tax law reveals a host of definitions and "imperfections" that are potential sources of additional inconsistencies and ambiguities: the distinction between capital and ordinary treatment, the "double taxation" of corporate income, the asymmetry in treatment between gains and losses, nonlinear (e.g., progressive) rate structures, and the body of source rules that deal with foreign taxpayers. In addition, the difficulties for the spanning method revealed by examining a simple economy as in the sections above would persist or intensify in a more realistic model with more than five periods and more than two states.[69]

The positive result that the spanning method can produce a consistent and universal tax system in the face of disparate tax treatments for different asset types is nonetheless interesting. Even substantial reform is likely to result in a "hybrid" tax system that does not apply the same treatment to all assets. It is useful to know that such systems may accommodate theoretically distinct approaches for different assets and still be consistent and universal. Furthermore, as discussed in the next Part, several approaches currently advocated for dealing with new financial products can only be successful if they reduce to the spanning method.

II. BEYOND THE SPANNING METHOD: A GENERAL FRAMEWORK

As noted above,[70] the prevailing view seems to be that global pattern taxation is the only comprehensive solution to the problem of taxing new financial products. If a particular generic tax treatment (such as accretion taxation or cash flow taxation) applies to all existing financial instruments, the government can simply apply this same treatment to any "new" financial product. While difficulties in implementation that require approximations and compromises might arise,[71] the approach would be clear and coherent.

---

68. *See* text accompanying notes 59-60 *supra*.

69. The "real world" includes at least five outcomes and two time periods. Although increasing the number of states would increase the freedom to specify tax treatments, direct inconsistencies would not vanish. In addition, an increase in the number of states would make computing the proper tax treatment for a "new" asset more complex and probably would lead to worse tax aesthetics.

70. *See* text accompanying notes 13-16 *supra*.

71. *See, e.g.*, Jeff Strnad, *Periodicity and Accretion Taxation: Norms and Implementation*, 99 YALE L. J. 1817, 1891-99 (1990) (discussing the implementation of accretion tax when asset price paths are unknown).

The previous Part shows that by using the spanning method it is possible to design a consistent and universal system for taxing new financial products that does not require prescribing a comprehensive treatment for all transactions. The above analysis does not indicate, however, whether the spanning method is the only approach with that property. In fact, spanning method approaches are only a subset of the set of all successful approaches. Some alternative approaches are less restrictive but still yield consistency and universality.

The major tasks of this part are to provide a logical taxonomy of theoretical approaches and then to relate this taxonomy to practical approaches such as bifurcation and integration that are discussed in the literature. To accomplish these tasks it is necessary to develop several refinements of the concept of consistency.

Section A shows that all methods that are consistent and universal can be expressed as integration schemes. Thus, the class of successful bifurcation and local pattern approaches is a subset of the class of successful integration schemes. Section B introduces the property of linearity, a property that implies consistency, and states two principal results. First, any linear and universal tax system reduces to the spanning method. Second, any consistent and universal bifurcation approach must be linear. As a result, bifurcation approaches that are consistent and universal are equivalent to the spanning method. In addition, local pattern taxation will be consistent and universal only if it reduces either to the spanning method or to an integration scheme. Section C explores integration schemes. These schemes can achieve a consistent and universal tax system even when the system is not required to be linear. This trait is important since modern tax systems tend to have significant nonlinearities. Section C also develops the concept of continuity, a condition that is weaker than linearity but stronger than consistency, and shows that all continuous and universal tax systems reduce to an integration scheme. In contrast, the spanning method will only succeed if the tax system is linear as well as continuous and universal. Continuity is an important condition because in its absence all of the undesirable effects that follow from inconsistency would be present.

A. Entire Integration and Consistency

Combining a taxpayer's portfolio into one single position and then associating a tax treatment to that position is a form of integration, called ''entire integration.''[72] Since any universal and consistent tax system associates to each total position a unique tax treatment, the following principle is true:

> *The Entire Integration Principle*: Any consistent and universal tax system is equivalent to an entire integration approach.

Some integration approaches may aggregate only certain groups of instruments but not the taxpayer's complete portfolio. Under the Entire Integration Principle, however, any consistent and universal integration method can be treated as if it did integrate the whole portfolio, since each aggregate position must result in the same tax treatments no matter which combination of components led to the position. Furthermore, any method that achieves consistency and universality, including the spanning method, also can be expressed as an entire integration scheme. Thus, all consistent and universal bifurcation methods and all consistent and universal local pattern approaches belong to the set of consistent and universal integration approaches.

The spanning method is one particular method of bifurcation that assures consistency and universality. The Entire Integration Principle indicates that the class of consistent and universal tax systems are precisely those that are equivalent to an entire integration scheme. Do all such systems *also* reduce to the spanning method? The answer is that they do not. In the next section we show that the class of universal tax systems that reduce to the spanning method is

---

72. Mathematically inclined readers will realize that there is an easy way to describe any entire integration scheme. It is a function that maps the vector space of portfolios into a vector space of tax outcomes.

It is easy to see that the set of all portfolios, as well as the set of all portfolio returns (see note 40 *supra*), form a vector space. First, choose a particular minimal spanning set. Each portfolio is a unique combination of assets in this set. The amount of each asset in the portfolio are the ''coordinates'' of the portfolio in the vector space. In the text example, this vector space is a five-dimensional Euclidean space, the coordinates for each dimension corresponding to the amount of a particular minimal spanning set asset in portfolios located in the space.

The space of tax outcomes consists of vectors specifying how much tax will be paid at the end of each period in each state of the world. In the text example, these vectors will be ten-dimensional, representing the taxes paid at the end of years one and two in each of the five states of the world. Of course, the set of all possible tax outcomes in the example will be isomorphic to a space with fewer than ten dimensions. The taxes paid at the end of the first year are not contingent on the state of the world. *See* text accompanying notes 36-37 *supra*.

If we denote the space of portfolios as ''P'' and the space of tax outcomes as ''T,'' then an entire integration scheme is a function θ that maps P into T. Requiring that θ be a function (rather than a correspondence) captures the feature that an entire integration scheme specifies a ***unique*** set of tax outcomes for each portfolio.

precisely the class of all such systems that are also *linear*. Some entire integration schemes are not linear and thus do not reduce to the spanning method.

B. Linearity, Bifurcation and the Spanning Method

A tax system is *linear* if the tax treatment of an asset or portfolio is the sum of the tax treatments of the components that make up the asset or portfolio.[73] Linearity requires that the tax treatment of any portfolio must remain the same regardless of the manner in which the portfolio is divided into particular assets. As a result, linearity implies consistency. Section C shows that there are consistent tax systems which are not linear. Thus, linearity is a stronger condition than consistency.

The class of tax systems that are both linear and universal is exactly the class that can be implemented by the spanning method:

> *Spanning Method Principle*: Any universal and linear tax system is equivalent to applying the spanning method using any minimal spanning set. Conversely, the spanning method will result in universality and linearity under any choice for the minimal spanning set.

Proving this principle is straightforward. Suppose a tax system is universal and linear, and choose any minimal spanning set. Since the system is universal, it prescribes a tax treatment for every asset, including each asset in the chosen minimal spanning set. Since the system is linear and since it is possible to express any cash flow pattern as a combination of assets from the minimal spanning set, the tax treatment of the spanning set elements uniquely determines the tax treatment for all cash flow patterns and assets. Thus, all linear and universal tax systems reduce to the spanning method. It is easy to show the converse: The spanning method guarantees that the tax system is linear.[74]

---

73. It is easy to define linearity mathematically using the framework of note 72 *supra*. In this framework a function θ that maps portfolios to tax outcomes represents the tax system. That system will be linear if for any two portfolios x and y and for any two numbers q and r (representing ways of mixing the portfolios):

$$q\theta(x) + r\theta(y) = \theta(qx) + \theta(ry) = \theta(qx + ry).$$

74. The converse follows from the assumption that the rule for computing the tax on combinations of minimal spanning set asset under the spanning method is linear. *See* text accompanying notes 45-47 *supra*. A portfolio is a weighted sum of a collection of assets, and each asset, in turn, is a unique canonical combination of assets from the minimal spanning set. The portfolio is also a unique canonical combination of spanning set assets. Since each cash flow pattern has a unique representation in terms of minimal spanning set assets, the canonical combination for the portfolio must equal the sum of the canonical combinations for the assets

There is a natural connection between the concept of linearity and bifurcation methods. ''Pure'' bifurcation approaches permit any decomposition of a particular portfolio for the purpose of computing taxes. An example is the bifurcation scheme in the 1991 version of the Treasury Department's proposed regulations for contingent debt. This scheme does not specify a particular method of decomposing contingent debt into pieces in order to compute the tax treatment of the whole.[75] Critics of ''bifurcation'' in general, as well as critics of the specific proposal in Treasury's 1991 proposed regulations, emphasize the possibility that different decompositions may lead to different tax treatments.[76] Where arbitrary decompositions are permitted under a pure bifurcation method, this problem will exist unless the tax system is linear. Linearity is a necessary condition for pure bifurcation to work since linearity asserts that the tax treatment of the whole is the sum of the tax treatments of the parts. But linearity is also a sufficient condition since linearity implies consistency. In a linear system, different decompositions cannot lead to different aggregate tax consequences.

Conversely, pure bifurcation (such as the scheme envisioned in the 1991 contingent debt regulations) will not work in a nonlinear tax system. In such a tax system there will be at least one portfolio and at least one decomposition of that portfolio under which the sum of the tax treatments of the parts does not add up to the tax treatment of the whole. To be consistent, the tax system cannot permit an arbitrary decomposition of any such portfolio for tax purposes. Instead, the tax system must assign a specific tax treatment to the portfolio that is independent of the treatment that would emerge from summing the taxes under particular decompositions. This feature is the hallmark of integration: The taxpayer does not have the freedom to characterize a transaction or portfolio position according to the tax treatments of the pieces but must apply a single specified tax treatment to the whole. To avoid the requirement of linearity but still retain consistency, an element of integration must be mixed with any bifurcation approach.[77]

---

comprising the portfolio. *See* text accompanying notes 40-41 *supra* (Unique Representation Principle). Because the rule for computing the tax on combinations of minimal spanning set elements is linear, the tax on a combination from that set must equal the sum of the taxes on the components in any linear decomposition of that combination. As a result, the tax on any portfolio will equal the sum of the taxes on the assets that comprise the portfolio.

75. *See* note 7 *supra* and accompanying text.

76. *See* note 14 *supra* and accompanying text.

77. For example, one might bifurcate according to one minimal spanning set for one type of transaction and according to a second minimal spanning set for another type. *See* note 45 *supra* and accompanying text. This approach would result in a tax system that is not linear if the tax treatments of certain transactions depend on which spanning set applies. To ensure consistency,

There is not much new to say in considering local pattern taxation. As demonstrated above, every consistent and universal tax approach is equivalent to some integration scheme.[78] Moreover, it is possible to implement any linear and universal system using a bifurcation scheme based on the spanning method.[79] Departures from linearity require that the tax system include an integration element in order to maintain consistency.[80] Whether a system that employs local pattern taxation is linear or nonlinear, therefore, this ostensibly "alternative" approach is actually extraneous, at least for analytic purposes.

C. *Integration Revisited: Nonlinear Tax Systems and Continuity*

Integration by nature ignores the composition of portfolios in favor of their aggregate. As a result, under integration, a tax system can be consistent without being linear. The sum of the tax treatments of assets in a portfolio need not be the same as the tax treatment of the portfolio. Since consistency is a weaker

---

the tax system would have to specify the treatment of such transactions, independent of the treatments that might arise from some possible decompositions. An element of integration would therefore be present.

The spanning method avoids the need to use integration by requiring that *one* particular minimal spanning set be chosen for decomposing *all* instruments. *See* text accompanying note 45 *supra.*

78. *See* text accompanying note 72 *supra* (Entire Integration Principle).

79. *See* text accompanying note 73-74 *supra* (Spanning Method Principle).

80. *Cf.* text accompanying notes 76-77 *supra* (integration element must accompany bifurcation method when tax system is nonlinear).

Introducing a local pattern may *create* nonlinearities. More specifically, if a particular portfolio of "old" instruments add up to a "new" instrument subject to the local pattern, nonlinearity (and inconsistency) will result unless the sum of the tax treatments of the old instruments in the portfolio happen to add up to a tax treatment that conforms to the local pattern. Avoiding inconsistency in this situation by ignoring the tax treatments of old instruments when they are combined to produce a new instrument amounts to an integration approach.

The ABA Report on the tax treatment of contingent debt advocates an integration strategy as part of a local pattern approach. In particular, the Report calls for integrating the contingent portion of such debt and for taxing that part as a unit under a set of generic (local pattern) rules rather than attempting to decompose it into "component parts" for tax purposes. *See ABA Report*, *supra* note 3, at 1189 (recommendations (13) and (15)-(20)), 1195. The stated rationale for this proposal is that bifurcation is a poor vehicle for achieving consistency because there is no unique way to decompose the contingent portion of the debt. *See id.*, at 1194-95. However, the report is also skeptical about whether the integration approach itself could succeed in achieving consistency. *See id.*, at 1195.

condition than linearity, successful integration approaches are a broader class than approaches based on the spanning method.

Since every consistent and universal tax system reduces to an entire integration scheme,[81] classifying consistent and universal tax systems is a matter of classifying entire integration approaches. One part of this task is finished. From the Spanning Method Principle,[82] it is clear that the set of all linear and universal systems is precisely the set of all systems that reduce to the spanning method:

> *Linearity Property*: An entire integration scheme (*i.e.*, any consistent and universal tax system) is linear if and only if the scheme can be generated by the spanning method.

To further refine the classification of integration approaches, it is useful to introduce *continuity*, a concept that lies between linearity and consistency. An entire integration scheme is *continuous* if portfolios that are nearly identical have nearly identical tax treatments.[83] In particular, small changes in any portfolio will not cause a "jump" in the tax results.

Continuity is a stronger property than consistency. A tax system is consistent if there is a unique tax treatment for each cash flow pattern and if the law treats the long and short versions of each position in a symmetric way.[84] Continuity adds the requirement that the difference in tax treatment for any two positions must approach zero as the two positions converge.

---

81. *See* text accompanying notes 72 *supra* (Entire Integration Principle).

82. *See* text accompanying notes 73-74 *supra*.

83. More rigorously, a tax system is continuous if for any position and any positive number (no matter how small), it is possible to choose a range of portfolios surrounding the position such that the tax treatment of each portfolio in the range differs from the tax treatment of the position by less than the chosen number. For this statement to make sense mathematically, one must specify a norm for the vector space of positions and a norm for the vector space of tax outcomes. The Euclidean distance norm is one possible choice for both spaces: for any two vectors x and y in n-dimensional space, $|x - y|$, the norm of x - y, is the square root of the sum of the squared differences between the coordinates of x and y.

Using the Euclidean distance as a norm, the usual definition of continuity applies: An entire integration scheme θ that maps from the space of portfolios to the space of tax outcomes is continuous if for any position, p, and for any $\varepsilon > 0$, there is a $\delta > 0$ such that $|p - p^*| < \delta$ implies that $|\theta(p) - \theta(p^*)| < \varepsilon$.

84. *See* note 54 *supra*.

A policymaker who values consistency will value continuity for the same reasons. More importantly, the absence of continuity may vitiate many of the benefits of consistency. Consider, for example, a system that is consistent but discontinuous, with a ''jump'' in tax treatments at a particular point. It would then be possible to achieve ''approximate'' inconsistency by matching a long position very close to the jump point with a short position situated exactly at the point. The pre-tax net cash flow from these two positions can be made arbitrarily close to zero while leaving a significant net tax effect. Investors can therefore come arbitrarily close to achieving pure tax arbitrage.[85] Thus, a system that is consistent but discontinuous provides opportunities for tax manipulation that are almost identical to the manipulations possible in an inconsistent system.[86]

As mentioned above, continuity implies consistency and therefore is a stronger requirement to impose on a tax system.[87] However, continuity is a weaker requirement than linearity. Under fairly innocuous assumptions about the ''finiteness'' of taxes,[88] it is a mathematical fact that linearity implies

---

85. Suppose, for example, that the tax system is consistent but that there is a particular point where the tax treatment shifts from capital to ordinary. Suppose also that portfolios near this point produce gains in most states of the world. One could set up a long position slightly on the capital side of the discontinuity and a short position slightly on the ordinary side. By moving these positions closer and closer together toward the actual point of discontinuity, one could make the combined cash flow from the positions arbitrarily close to zero. At the same time, the tax treatments of the two positions will match capital gains on the long side against ordinary losses on the short side, thereby creating a net tax advantage (conversion of ordinary income to capital gain income) with virtually no net pre-tax cash flow. The tax advantage will not diminish as the positions move closer and closer together, so long as they remain on opposite sides of the discontinuity. The taxpayer will have achieved ''approximate tax arbitrage.'' *See* note 12 *supra* (using similar example to discuss ''pure tax arbitrage'' where a taxpayer can reduce pre-tax cash flow all the way to zero).

86. *See* note 85 *supra* (providing example).

Concern for continuity is evident in the tax policy literature. For example, the ABA Report on the treatment of contingent debt proposes to exclude instruments with *de minimis* contingent payments from the ambit of the contingent debt regulations. This measure would prevent issuers from being able to ''elect'' contingent debt treatment merely by including in the debt package a contingent payment that is nearly valueless. *See ABA Report*, *supra* note 3, at 1192. This type of borderline issue would not arise in a tax system that was continuous.

87. Continuity requires that each portfolio map to a unique tax treatment. *See* note 83 *supra* and accompanying text (definition of continuity).

88. The only assumption necessary is that the tax liability for any finite collection of assets will be finite. This assumption is met if there is a finite minimal spanning set such that one unit of each asset in the set incurs a fmite tax in every state of the world. Given that real world economies do not produce infinite returns, the idea that no security will lead to infinite taxes (in present value) is believable.

continuity.[89] In contrast, functions can be continuous but not linear.[90] Consequently, there are continuous and universal tax systems that cannot be generated by the spanning method. Since these same tax systems can be generated by integration schemes (at least by entire integration schemes), it is clear that integration is a more general method for achieving desirable tax results than any approach that reduces to the spanning method:

> *The Integration Dominance Principle*: Every continuous and universal tax system can be generated by an entire integration scheme. But a tax system must be linear as well as continuous and universal in order to be generated by the spanning method.

Because bifurcation methods tend to succeed only in linear tax systems, the Integration Dominance Principle validates the common intuition among practitioners that integration approaches have greater potential than bifurcation methods.[91]

This greater potential is of more than theoretical significance. The current U.S. tax system has many nonlinearities. Only integration methods will achieve continuity or consistency in such a system. One example of a nonlinearity in the system is the asymmetry in the treatment of losses and gains. Suppose that a start-up company can do project A or project B or both. Project A yields a $3000 profit or a $1000 loss with equal probability while project B yields $1000 with certainty. If the company does only project A and sustains a $1000 loss, the company cannot deduct the loss immediately but must carry the loss forward and use it against future gains. However, doing project A and project B simultaneously results in immediate use of the loss since the loss offsets the $1000 gain from project B.

---

89. This result will be familiar to readers with some background in real analysis. A linear transformation between vector spaces is continuous if it is "bounded." A transformation is "bounded" if all portfolios on the unit ball of the portfolio space have finite tax outcomes. Because all portfolios on the unit ball are weighted combinations of assets in some finite minimal spanning set and because the tax system is linear, the finiteness of the tax vectors corresponding to the minimal spanning set assets guarantees continuity.

90. For example, consider the function $f(x) = x^r$, where x takes values in the nonnegative real numbers, and r is some positive real number other than zero or one. For any such value of r, this function is continuous but not linear.

91. *See, e.g.*, Kau, *supra* note 3, at 1005-1007 (arguing for integration since it avoids the complexities inherent in bifurcation); *ABA Report*, *supra* note 3, at 1194-95 (advocating integration over bifurcation for the contingent part of debt).

These loss rules result in a nonlinearity. To see why, suppose that the applicable tax rate is 40 percent and that the company will not do any projects other than A or B. The taxes arising from doing project A and project B separately *do not* always add up to the tax arising from doing the projects together:

Taxes for Various Outcomes and Project Combinations (Given a $1000 Gain from Project B)		
	$3000 gain outcome for Project A	$1000 loss outcome for Project A
Project A alone	$1200	$0
Project B alone	$400	$400
Projects A and B together	$1600	$0

In the case where the outcome of Project A is a $1000 loss, the separate projects incur a sum of $400 in taxes, but doing the projects simultaneously results in zero tax liability.

The treatment of losses under current law is effectively part of an integration approach. That approach requires aggregating the income and deductions from all of a firm's projects and then applying a special tax rule when there is an aggregate loss. Because the ensuing tax system is nonlinear, the spanning method cannot generate the same result.

Despite the greater promise of integration methods, these methods are not a panacea for the current U.S. tax system. Any revision aimed at making the system consistent and universal would have to eliminate existing direct inconsistencies such as the ambiguous treatment of ''stocks'' and ''bonds'' described above.[92] This task would require fundamental reform, whether accomplished using integration approaches or otherwise.

92. *See* text accompanying notes 59-60 *supra*.

III. SUMMARY AND CONCLUSIONS

A. *An Overview of the Framework and its Application*

Part II developed a wide variety of results. This section provides an overview of the results and how they apply to the challenge posed by new financial products.

Before beginning the overview, it is profitable to begin with a point that at first glance may appear to be of limited relevance: There are many areas of tax law that do not explicitly deal with new financial products but that employ or could employ bifurcation and integration techniques.

Part II mentioned the integration aspects of the loss rules in current law. Another, more general, example of integration is the power of administrators and courts to apply the doctrine of "substance over form."[93] This doctrine permits the administrators or judges to ignore the individual "pieces" of a transaction and to tax the whole transaction according to their own view of the "substance."[94]

Some examples of bifurcation approaches in current law stem from the body of rules that apply to the sale of sole proprietorships and partnership interests. In the case of a sole proprietorship, complete bifurcation determines the tax due upon sale. Rather than treat the business as an integrated entity, the tax system

---

93. *See* BORIS I. BITTKER & LAWRENCE LOKKEN, FEDERAL TAXATION OF INCOME, ESTATES AND GIFTS at ¶ 4.3.3 (2d ed. 1989).

94. *See* Knetsch v. United States, 364 U.S. 361 (1960). In *Knetsch* the taxpayer combined an investment in annuities with a yield of 2.5 percent per year with offsetting debt carried at a 3.5 percent annual rate. By agreement, the taxpayer could increase the amount of debt each year by the accrued 2.5 percent increase in value for the annuity. The 3.5 percent interest payments on the loans resulted in a current deduction while the gain from the increased value of the annuities was both deferred and subject to favorable capital gains rates under the rules applicable at the time. Although the taxpayer incurred negative cash flows throughout the transaction (even assuming continual borrowing of the 2.5 percent annual accrual to the annuity contracts), the substantial tax advantages made the whole transaction quite profitable.

Using a "substance versus form" approach, however, the Supreme Court denied the interest deduction for the borrowing. *Id.* at 366 (stating "it is patent that there was nothing of substance to be realized by Knetsch from this transaction beyond a deduction"). The Court implicitly used an integration approach since it considered the nature of the whole transaction rather than allowing the taxpayer to treat each part as dictated by the then current law.

assigns a gain or loss, whether capital or ordinary, to each asset of the proprietorship as if sold individually.[95] In contrast, the sale or liquidation of a partnership interest triggers a partial bifurcation approach. To the extent of a partner's share in certain "hot" assets, principally appreciated inventory and accounts receivable, gain upon sale is ordinary income instead of capital gain.[96] This distinct treatment applies only to some "ordinary" assets of the partnership, so that the bifurcation is incomplete.[97]

Since bifurcation and integration pervade current law as well as proposals to address new financial products, the results developed concerning bifurcation and integration in Part II apply to many issues in current law as well as to provisions aimed at specific financial innovations. But there is also a deeper interrelationship. If current law uses flawed bifurcation or integration schemes, specific flaws can become focal points for tax-motivated financial creativity. It is the failure of the tax system to be consistent, continuous or linear that underlies any situation where "choice" of tax treatments based on alternative ways of structuring transactions is a concern.

It is worth summarizing the role that these three principles play with special emphasis on the implications for "tax innovation." The three are logically related. Linearity, the requirement that the sum of the taxes on the components of a portfolio add up to the tax on the whole portfolio under every possible decomposition, is the strongest principle since it implies consistency and, under a reasonable "finiteness" assumption about taxes, continuity.[98] Continuity is the next strongest, being in essence a generalization of consistency. Continuity requires convergence of tax treatments as portfolios converge to a single portfolio, while consistency requires only that any single portfolio have a single, specified tax treatment.[99]

Bifurcation methods will succeed in general only if the tax system is linear.[100] The fact that current law has significant nonlinearities means that

---

95. *See* BITTKER & LOKKEN, *supra* note 93, at ¶51.9. The seminal case establishing this rule is Williams v. McGowan, 152 F.2d 570 (2nd Cir. 1945). *See id.*

96. *See* I.R.C. §751(a) (1994).

97. In addition, partners can structure liquidations, adjust inventory holdings prior to sale, or make other adjustments to avoid the operation of the rules. *See* 2 WILLIAM S. MCKEE, WILLIAM F. NELSON & ROBERT L. WHITMIRE, FEDERAL TAXATION OF PARTNERSHIPS AND PARTNERS, ¶ 21.01-21.06 (2d ed. 1990).

98. *See* text accompanying notes 88-89 *supra.*

99. *See* text accompanying notes 83-86 *supra.*

100. *See* text accompanying notes 75-77 *supra.*

policymakers should be very cautious about these methods. Practitioners are familiar with many of these nonlinearities and the opportunities that exist to exploit them when the tax law uses a bifurcation approach. Consider, for example, the bifurcation rule mentioned above that applies to sale of a partnership interest. The general rule is that sale or exchange of a partnership interest results in capital gain or capital loss.[101] However, partnership holdings are divided into categories and to the extent that the partnership has assets in certain categories, gain on sale may be characterized as ordinary gain.[102] One of these categories is ''substantially appreciated inventory,'' and this category includes all of the partnership's inventory *if* this inventory has appreciated by at least 20 percent (over its adjusted basis) *and* comprises more than 10 percent of the partnership's noncash assets.[103] The rule does not consider assets held by the partners outside of the partnership. Thus, the partners can contribute some of their noncash assets, such as short-term Treasury bills, to the partnership in order to ensure that inventories are less than 10% of the noncash assets.[104] When the partnership is sold, the partners will receive cash corresponding exactly to the assets that they have contributed and can reinvest in these assets. This set of maneuvers involves no actual change in the partners' portfolios. Instead of holding certain assets outside of the partnership, they (temporarily) hold the same assets as partners. The result of this repackaging is a reduction in tax liability when the partnership interests are sold.

New financial instruments provide new methods for dividing up cash flow patterns and thus more potential for exploiting nonlinearities. As it becomes easier and less expensive to invent new instruments, nonlinearities in the tax code that coexist with bifurcation approaches become a much more serious problem. Indeed, as mentioned above, some of these nonlinearities may be focal points for financial innovation. In addition, using bifurcation to address problems that arise from new financial products may fail, or even create new problems, in the face of nonlinearities.

Successful application of integration methods does not require linearity. However, continuity becomes a serious concern. If portfolios with nearly identical returns have widely divergent tax implications, a taxpayer can ''choose''

---

101. *See* I.R.C. §741 (1994).

102. *See* I.R.C. §751(a) (1994).

103. *See* I.R.C. §751(a) and (d)(1) (1994).

104. *See* MCKEE, NELSON & WHITMIRE, *supra* note 97, at ¶16.04[2] (short-term, cash-like investments are included for purposes of applying the ten percent rule while cash itself is not).

a radically different tax treatment by very slight portfolio changes,[105] or can engage in approximate tax arbitrage by being short in one of the two portfolios and long in the other.[106] An advanced financial engineering industry that can produce new instruments at low cost exacerbates these problems since the availability of new instruments makes finer gradations in portfolio choice possible.[107]

There are many familiar examples of discontinuities in current law. One need only look for sharp borderlines that delineate abrupt changes in the aggregate tax treatment of an asset or transaction. The borderlines between capital assets and ordinary assets, between short-term and long-term gains, and between passive income and active income are a few major examples. Another more detailed example is the substantially appreciated inventory test just discussed. A small change in inventories or asset composition in a partnership can determine whether the partnership passes the test and thus have very significant tax consequences for the partners.[108]

Existing discontinuities are focal points for tax-motivated financial innovation since new financial instruments permit a closer approach to the borderlines

---

105. One can extend the continuity concept by requiring that tax consequences not change "too quickly" with any particular portfolio shift. Continuity guarantees that there are no sudden "jumps" in tax treatment as investors vary the composition of their portfolios. However, one can imagine a sequence of continuous tax rules that involve increasingly drastic (but continuous) shifts in tax treatment near a given point in portfolio space and that will converge to a jump at that point in the limit. Continuous tax rules that dictate large shifts in tax treatment for small portfolio shifts can come arbitrarily close to tax rules with a discontinuity.

106. *See* note 85 *supra* (example of approximate tax arbitrage).

107. Short and long positions that are close to being exact offsets may not always indicate attempts at manipulation. Many of these matched positions may signify attempts to hedge business risk where only approximate hedging is available at an acceptable cost. For example, a company may have issued bonds with a payment at maturity that depends on the price 3-month Treasury bills at that time. The company might wish to hedge these bonds with Treasury bill futures. However, the most heavily traded Treasury bill futures contracts have standard expiration dates in March, June, September, and December. *See* DARRELL DUFFIE, FUTURES MARKETS 346, 350 and 353 (1989). The bonds may mature in some other month so that there is an imperfect match up between the publicly-traded Treasury bill futures and the Treasury bill price risk inherent in the bonds. Contracting privately for a Treasury bill futures contract with precisely the right expiration date might be so costly that the company's best strategy is settle for an imperfect hedge. Hedging business risks can serve socially valuable purposes. *See* note 111 *infra*.

108. The test causes a discontinuity because passing over a particular numerical borderline abruptly changes the treatment of an aggregate of assets, in this case, inventories. Given this discontinuity, it is not surprising that the test also results in a nonlinearity. *See* text accompanying notes 100-104 *supra*. Linearity is a stronger condition than continuity, so that a discontinuous tax system cannot be linear. *See* text accompanying notes 87-90 *supra*.

between tax treatments. A closer approach means that investors will not need to alter the preferred pre-tax qualities of their portfolios as much in order to achieve desirable tax outcomes and that ''approximate'' tax arbitrage will approach pure tax arbitrage more closely. Thus, when integration schemes are on the agenda, it is important to consider how these schemes exacerbate or alleviate the problems caused by discontinuities.

The serious administrative costs of discontinuities and nonlinearities are apparent from history. These flaws induce major struggles between taxpayers and the govermnent, often accompanied by new legislation attempting to curtail abuse. Thus, for example, the distinction between capital and ordinary assets, the distinction between short-term and long-term gains and the substantially appreciated inventory rule have each resulted in massive on-going administrative problems.[109]

B. Tax Policy Implications

If current trends are any indication, the variety and heterogeneity of available financial products will continue to increase rapidly.[110] Such growth threatens to create greater and more frequent problems for a tax system already riddled with inconsistencies and discontinuities. Hence two profound challenges face those who create, administer, and interpret the tax code: First, how can the legislative branch of govermnent shape the tax code in directions conducive to a ''good'' tax system? And second, how should administrators and courts respond to the dynamic tax enviromnent created by financial innovation given that they cannot alter the major choices made by the legislature?

The above analysis provides a theoretical structure useful for the debate concerning these ongoing questions. Each of the popular approaches for taxing new financial products fits into a larger framework.

One way to achieve consistency and universality is to construct a tax system with a single ''global pattern'' of taxation such as cash flow taxation or accretion

---

109. *See, e.g.*, BITTKER & LOKKEN, *supra* note 93, at ¶51 and ¶53 (overview of rules defining ''capital assets'' and holding period rules); MCKEE, NELSON & WHITMIRE, *supra* note 97, at 16.04[2] (manipulation possibilities currently available under substantially appreciated inventory rules and potential government responses); LIND, SCHWARZ, LATHROPE & ROSENBERG, FUNDAMENTALS OF PARTNERSHIP TAXATION, 243-44 (Foundation Press 3rd ed 1992) (elimination by statute of second-tier partnership strategy for exploiting substantially appreciated inventory rules).

110. *See* note 1 *supra*.

taxation. But this extreme degree of homogeneity is not necessary. A linear tax system relying on the spanning method can harbor radically different treatments for different financial instruments. Even linearity is not required. The most powerful and general approaches relying on integration methods can operate in a nonlinear environment. In fact, any consistent and universal set of rules is equivalent to an integration approach. However, an integration scheme must be continuous, rather than merely consistent, or manipulation through approximate "tax arbitrage" and other devices will be possible.

The current U.S. tax system contains not only nonlinearities and discontinuities but also direct inconsistencies. For example, the tax treatment of equity and debt differ substantially, but there is considerable overlap in these categories. As a result, to a significant extent, financial engineers can package the same set of cash flows to be equity or debt for tax purposes.

Repairing the major discontinuities and inconsistencies in current law is a task that would require fundamental reform. These discontinuities and inconsistencies arise from the debt/equity distinction, the distinction between capital assets and ordinary assets, the differential treatment of gains and losses by holding period and other aspects of current law that are central to the statutory scheme. Changing these aspects is only possible at the legislative level. Addressing major inconsistencies and discontinuities at that level would go a long way toward relieving the pressure arising from financial innovation. As discussed in the previous section, these flaws in the tax system tend to act as focal points that stimulate the development of tax-motivated new financial products. In addition, if administrators must take these inconsistencies and discontinuities as given and immutable, not even the most powerful integration techniques will succeed at making the tax system consistent or continuous.

Even if the legislative branch implements some major changes, it is likely that some significant inconsistencies and discontinuities will remain. This fact and the darker possibility that no major reform will occur create some difficult choices for authorities such as the Treasury Department and the courts. These authorities are necessarily limited to prescribing "second best" solutions since no set of Treasury Regulations or cases can guarantee universality and consistency or continuity in the face of major inconsistencies and discontinuities that must be taken as given. Perhaps the only viable alternative for these authorities in addressing financial innovation is the traditional one of analyzing the normative stakes for each type of transaction and then crafting a detailed response.[111]

---

111. The discussion in the literature concerning the tax treatment of contingent debt provides a good example of this type of analysis. Issuers of contingent debt often hedge the contingencies by purchasing or issuing other financial instruments. This hedging allows the issuer to shift the risk from the contingent portion of the debt to others, thereby facilitating a transaction (issuance of the

Since these stakes differ by type of transaction, comprehensive rules will not

contingent debt) that can reduce the issuer's borrowing costs while simultaneously making capital markets more efficient by providing an instrument that investors want. *See* Kleinbard, *supra* note 3, at 954. To ensure that asymmetric tax treatment of the hedge position and the contingent debt will not discourage such transactions, several commentators have advocated integrating the two and treating the entire package as debt. *See* note 9 *supra* and accompanying text.

Not all corporate hedging transactions are socially valuable or are good for shareholders. *See* Hu, *Shareholder Welfare*, *supra* note 2, at 1306-1309 (hedging by public corporations can be wasteful when it is costly but merely eliminates risk that shareholders have already eliminated by holding diversified portfolios). But business hedging can serve socially valuable purposes. For example, investors may not be able to distinguish poor business performance due to market fluctuations from poor performance due to poor management. By hedging business profits, managers can eliminate the market fluctuation effect and thus provide clearer information to these investors. It would be unfortunate if the tax system discouraged financial innovation that permits socially valuable hedging. However, the private and social costs of any particular financial innovation may be quite subtle. *See* Hu, *Misunderstood Derivatives*, *supra* note 1, at 1465-67, 1465 n.31; Hu, *Regulatory Paradigm*, *supra* note 2, at 364-66. As a result, choosing the best tax treatment for financial innovation involves some fairly difficult issues pertaining to capital markets. It is not clear as an a priori matter which financial innovations are worth encouraging and which are not. In addition, it is not clear whether the tax laws are the appropriate vehicle to impose a subsidy or a burden.

The picture is even more complicated because there are also a rich set of tax motivations for hedging. A firm may hedge to reduce fluctuations in taxable income and thereby reduce the possibility of a delay, due to the loss carryforward rules, in the tax deduction for current losses. Hedging also allows a firm to borrow more. It appears that the tax code currently discriminates against firms with low borrowing capacity, and crafting tax rules that favor hedging (or at least do not hinder it) may correct this discrimination. *See* Michael S. Knoll, *Taxing Prometheus: How the Corporate Interest Deduction Discourages Innovation and Risk-Taking*, at 20-25 (forthcoming *Villanova L. Rev.*, March 1994). Clearly, these considerations suggest a whole host of "second best" tax policy issues. The policymaker's attitudes toward the corporate interest deduction, the treatment of losses and even corporate tax integration are relevant.

always be desirable.[112] Loose ends in the form of inconsistencies, discontinuities or lack of universality will be inevitable.

112. For example, it may be optimal to apply differing treatments to financial innovations connected with different types of hedging transactions. *See* note 111 *supra*. In any event, the accumulation of regulations, each responding to particular instruments or special problems, may be unavoidable. Proposals that call for "common law development" or "a continuing dialogue" between Treasury, practitioners and taxpayers are not necessarily either soft-headed or a cop out. *See* note 16 *supra* and accompanying text. The gradualist theme in these proposals echo the more elaborate and general argument for an incremental, process-based approach to the regulation of new financial products made by Professor Henry Hu. *See* Hu, *Misunderstood Derivatives*, *supra* note 1, at 1495-96, 1513; Hu, *Regulatory Paradigm*, *supra* note 2, at 413-18, 435. Professor Hu stresses the informational disadvantages faced by regulators who must integrate new financial products into regulatory structures developed in response to existing financial instruments. Hu, *Misunderstood Derivatives*, at 1463, 1495-1508; Hu, *Regulatory Paradigm*, at 405-16. Although Professor Hu deals primarily with financial regulation such as the capital adequacy rules that apply to banks, many of his arguments have obvious parallels for the design and administration of the tax laws in the face of financial innovation.